

NBER WORKING PAPER SERIES

TRADE IN UNIVERSITY TRAINING:
CROSS-STATE VARIATION IN THE PRODUCTION
AND USE OF COLLEGE-EDUCATED LABOR

John Bound
Jeffrey Groen
Gábor Kézdi
Sarah Turner

Working Paper 8555
<http://www.nber.org/papers/w8555>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2001

We have benefited from the comments of seminar participants at the NBER Summer Institute, the Upjohn Institute, the University of Michigan, the University of Virginia, the University of Kentucky, the University of Toronto, Cornell University, and the Bureau of Labor Statistics. We are grateful for conversations and detailed comments from Charles Brown, Mike Chernew, Julie Cullen, James Hearn, Gordon Hanson, Larry Kahn, Abigail Payne, and Gary Solon. We thank the Russell Sage Foundation, the Andrew W. Mellon Foundation, the Bankard Fund for Political Economy, and the Population Studies Center at the University of Michigan for research support. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2001 by John Bound, Jeffrey Groen, Gábor Kézdi and Sarah Turner. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Trade in University Training: Cross-State Variation in the
Production and Use of College-Educated Labor
John Bound, Jeffrey Groen, Gábor Kézdi and Sarah Turner
NBER Working Paper No. 8555
October 2001
JEL No. J0, I2

ABSTRACT

The main question addressed in this analysis is how the production of undergraduate and graduate education at the state level affects the local stock of university-educated workers. The potential mobility of highly skilled workers implies that the number of college students graduating in an area need not affect the number of college graduates living in the area. However, the production of relatively large numbers of college and university graduates in an area may lead to increases in the employment of university-trained manpower if local industries expand production of goods that use college-educated workers intensively. Using data from the U.S., we find a modest link between the production and use of BA degree recipients; states awarding relatively large numbers of BA degrees in each cohort also have somewhat higher concentrations of college-educated workers.

John Bound
University of Michigan and NBER

Jeffrey Groen
University of Michigan

Gábor Kézdi
University of Michigan

Sarah Turner
University of Virginia and NBER

In the United States, higher education draws heavily on the resources of state and local governments through direct subsidies and indirect subsidies in the form of exemption from taxation. A rationale often given for why states invest in the education of their residents is that states enjoy some of the returns from such investments – the more highly educated a workforce, the more productive it is. What is more, highly educated workforces may contribute to regional economic growth by attracting new business. In fact, there is increasing evidence that the overall skill level of an area’s workforce has fundamental effects on the local economy.¹ However, given the mobility of the labor force in general (Long, 1988; Bartik, 1991; Blanchard and Katz, 1992) and of college-educated labor in particular (Long, 1988; Bound and Holzer, 2000), there may be little correspondence between the number of college students graduating in a state and the number of college graduates living in the area. The question addressed in this analysis is whether the production of higher education in a state affects the local stock of human capital in a state.

Understanding the factors contributing to differences in the level of collegiate attainment across states is key to assessing the return to state subsidies for higher education. At issue is how policies affecting the “supply side” or the production of college-educated workers compare to other incentives affecting the location choice of college-educated workers. Framing this analysis at the state level reflects the observation

¹ Cities with well-educated workforces tend to grow faster than do cities with less well-educated workforces, with such differences persisting over time (Glaeser, Scheinkman, and Shleifer, 1995; Glendon, 1998). Moreover, wages of both well- and less-well-educated workers tend to be positively associated with the educational attainment of a city’s workforce (Rauch, 1993; Moretti, 1999). The evidence on whether this association is causal (Moretti, 1999; Acemoglu and Angrist, 2000).

that it is state policymakers who determine the level of institutional subsidy for higher education and the associated tuition rates at public colleges and universities.

Our work is also relevant for understanding the nature of the adjustments that occur in local area economies in response to supply shifts. Labor economists have typically emphasized the importance of migration as the means by which local areas respond to supply shocks (e.g., Borjas, Freeman and Katz, 1997). However, more recently researchers have emphasized the potential importance of changes in output mix in the adjustment of local economies to supply shocks (e.g., Hanson and Slaughter, forthcoming). As far as we know, no one has tried to quantify the relative importance of these two factors.

The first section of the paper outlines a model framing the relationship between changes in collegiate degrees conferred in a state (flows) and the supply of college-educated workers in an area (stocks). The second section outlines our empirical strategy and describes the data available to analyze this relationship. The third section describes the level and variation in the flow and stock measures across states and over time and then presents estimates of the relationship between flows and stocks over the long run and in response to transitory shocks. Because measurement error introduced by using institutional data on annual flows to approximate cohort flows is likely to introduce a non-classical errors-in-variables problem, we pay particular attention to the consequences for estimated parameters and use simulations to gauge the direction and magnitude of the potential bias.

A central finding of this paper is that the relative flow of degrees conferred within a state has a modest effect on the relative stock of university-educated workers within the

state and, as such, states have only limited capacity to influence the human capital levels in their workforces by investing in higher education degree outputs. For university graduates likely to be employed in sectors producing goods and services traded across states, we expect a stronger link between degree production and employment of college-educated workers than in sectors where goods and services are both produced and consumed locally. For BA degree recipients, who tend to work in the traded goods sectors of state economies, we find some positive association between the production and use of college-educated labor. In this general case, the presence of college graduates in a state works, to some extent, to attract industries that are intensive in such labor. A contrasting example is the case of MD degree recipients where there is virtually no relationship between production and use. Although the production of MD degrees tends to be concentrated in large, densely populated states, MDs disperse across the country after degree completion. Because MDs are concentrated in the non-traded sector of state economies, there is relatively little capacity for the presence of a large number of medical degree recipients in a state to attract medical industries to an area.

Section 1: Conceptual Framework

To aid in the interpretation of our estimates we develop a simple conceptual model. The purpose of the model is twofold. First, the model places the state-level degree outcomes of higher education in the context of supply and demand in the labor market,

thereby guiding the economic interpretation of our estimates.² The theoretical model also helps to frame the sources of bias likely to arise in estimating the causal effect of flows on stocks. This is important because credible instruments for our flow variable are not available.

Of primary interest is the effect of an exogenous change in the relative flow of college graduates (the share of people who graduate from college) on the stock of college-educated labor (the share of the workforce with college degrees) in an area. Figure 1 summarizes the intuition of the model, capturing the salient dimensions of the relative flow of college graduates, labor supply for college-educated workers relative to those at the high school level, and labor demand (Appendix 2 provides a formal presentation of the model). Focusing first on the top panel, the horizontal axis represents the relative supply of college-educated labor within the state, while the vertical axis represents wages for college-educated labor relative to high school-educated labor within the state. The F curve represents the relative flow of college-educated labor in the state arising from college graduation within the state. Without post-college migration, this would be the relative supply of college-educated labor to the state. The S curve incorporates migration. Under infinitely elastic migration, S would be horizontal at the national wage ratio. Panel 1 shows the case of imperfect but nonzero mobility, which gives a more elastic S curve than the F curve. The two curves cross at the wage level for which there is no net migration. For wages above this point there is net immigration of college-educated labor

² The model we present captures changes within a state in a partial equilibrium context. In this regard, wages outside the state are assumed to be exogenous and not affected by migration. We have confirmed that the qualitative implications of the model are robust to the partial equilibrium assumptions using a parameterized general equilibrium model of two equally large states. (These results are available from the authors.)

and S lies to the right of F ; for wages below this level there is net emigration of college-educated labor and S lies to the left of F .

D represents the relative long-run demand schedule for college-educated labor within the state. Shifts in the relative supply of college-educated labor in a state lead to adjustments in production, with shifts toward industries and technologies intensive in college-educated workers when this factor is relatively plentiful. For this reason, within state, we expect relative demand to be quite elastic. Indeed, in the textbook Heckscher-Ohlin model, relative demand curves are horizontal. In the top panel of Figure 1, the initial equilibrium occurs at point A. In the case represented, the state is a net importer of college-educated labor: equilibrium flows are smaller than stocks. We expect some states to be net exporters and others to be net importers.

An exogenous increase in the share of individuals graduating from college in the state is indicated as a shift in the schedule of flows in the top panel of Figure 1 from F to F' . The shift in F induces a shift in the net relative supply of college-educated labor in the state from S to S' , and the equilibrium shifts from point A to point B. The shift in S is likely to be somewhat smaller than the shift in F if, at the going wage rates, college graduates are more likely than high school graduates to migrate out of state. At the same time, we have drawn the curves under the simplifying assumption that the shift in F (and S) does not induce a shift in D – that is, assuming an increase in the relative flow of college graduates in the state does not directly affect the demand for college-educated labor.³ With the shift in the schedule of college graduates relative to high school

³ Such direct external effects would reflect technological complementarity between the production and use of college-educated labor. Adjustments associated with the inclusion of these effects are addressed in Appendix 2.

graduates from F to F' and the shift in the schedule of the relative supply of college-educated labor (the stock) from S to S' , equilibrium flows will shift from f to f' while equilibrium stocks will shift from s to s' . What we are interested in estimating is how a shift in relative flows (the shift from f to f') affects the change in the relative stock of college-educated workers in a state (measured by the adjustment in stocks from s to s'), in equilibrium.

The change in stocks over the change in flows (the ratio of $s'-s$ to $f'-f$ in Figure 1) is between zero and one and is a function of the demand and supply elasticities. More mobility dampens the effect of flows on stocks. At one extreme, no mobility leads to a one-to-one mapping between changes in flows and changes in stocks. At the other extreme, frictionless mobility leads to a zero effect of flows on stocks. In contrast, the larger the within-state elasticity of demand, the larger will be the effect of changes in flows on stocks. In the simple case depicted in Figure 1, the ratio of $s'-s$ to $f'-f$ will equal the ratio of the demand elasticity to the sum of the demand and migration elasticities⁴ (see Appendix 2 for details).

Although we have outlined the model in terms of the within-state market for college-educated labor, a similar framework can be used to analyze the labor market for more specialized kinds of skilled labor. We expect the parameters of the model to differ across markets for different types of skills. In particular, for types of skilled labor that work exclusively or almost exclusively in the non-traded sector of the economy (e.g., physicians, nurses, school teachers), we expect the within-state demand curves to be relatively steep. In this case, we also expect the relationship between the production and

⁴ In Figure 1, the demand elasticity is represented by the slope of the D curve, while the migration

use of the type of skilled labor in question to be small.⁵

We are interested in determining the effect of *exogenous* differences in flows on stocks. Across states, this variation can be thought of as due to states' "comparative advantage" in the production of higher education. However, substantial variation across states in local labor market conditions may confound the measurement of the causal effect of variation in flows on the long-term concentration of college-educated labor in a state. To see this, consider the extreme case where all of the variation in flows is determined by a change in demand. This case is illustrated in the bottom panel of Figure 1. The equilibrium in flows, stocks, and wages represented by point A is identical to the top panel. A shift in the demand function is represented by D' , which leads to movements along the flow and supply functions, with the new equilibrium established at point C. While demand-side variation also leads to a positive association between flows and stocks, the magnitude is quite different than in the initial case, with the change in stocks now exceeding the change in flows. What is more, the within-state equilibrium wage level rises rather than falling, as w' is greater than w in the bottom panel, while w' is less than w in the top panel. More generally, exogenous variation in the demand for college-educated labor will tend to bias the parameter estimate upwards, with the magnitude of this bias depending on the relative magnitudes of the exogenous cross-state variation in supply of college-educated labor and demand for college-educated labor (see Appendix 2

elasticity is represented by the difference in slopes between the S and F curves.

⁵ In our empirical work, we focus on medical doctors because the AMA maintains good data on stocks and flows. It has often been argued that physicians are capable of inducing a demand for their services. At the extreme, this would imply that the demand curve for physicians is perfectly elastic. However, the empirical evidence on this point (see, for example, Newhouse et. al., 1982) suggests that the location-specific demand for physicians is quite inelastic. Furthermore, in cases like medicine where the local supply elasticity is likely to be small, one might expect that employers and schools would work together to create institutions that would facilitate geographic mobility, further weakening the relationship

for an algebraic derivation of this point).

The interpretation of our results requires some understanding of the source of the cross-state variation in stocks and flows. Across states, variation in flows may reflect differences in the comparative advantage in producing college-educated labor, with differences derived from historical forces affecting the location choice of colleges more than a century ago, proximity to population centers, or willingness of voters to support higher education. Variation in demand for college-educated labor across states may occur because some states have a comparative advantage in the production of goods and services intensive in college-educated labor. The nation's political and financial capitals (Washington, DC and New York City) might be examples of this kind of phenomenon.

If states varied in their comparative advantage for production of higher education, but not in their use of college-educated labor, we would expect the states that produced the highest proportions of college graduates to be the states that used college-educated labor most intensively. Market forces would tend to induce those trained in high degree-production states to emigrate, but this phenomenon would not change rank orderings of the measures of flows and stocks by states. In this case we would expect to find a negative association between both stocks and flows and relative wages. In contrast, if states varied in their comparative advantage in the use, but not in the production of college-educated labor we would still expect to see a very high rank order correlation between states that produced a lot and states that used a lot of college-educated labor. In this case, we would expect to find a positive correlation between both the production and the use of college-educated labor and the relative wages of this group; however, causation

between flows and stocks.

would run from the labor market to the education market.

What we observe is that some of the states with high concentrations of college-educated workforces also produce a disproportionate share of college graduates, while others import college graduates. Likewise, some of the states that produce a disproportionate share of college graduates also have a disproportionate share in their work forces, while others export college graduates. This is consistent with the notion that there is cross-state variation in the comparative advantage in both the production and use of college graduates.

Section 2: Empirical Strategy and Data

Estimating Equations

In placing this model in an empirical context, we analyze the association between cumulative per-capita flows of degrees awarded for birth cohort g in state j and the per-capita stocks for the same cohort in the same state in some subsequent year t by estimating the following equation:

$$(1) \quad \ln \frac{Stock_{jgt}}{Population_{jgt}} = \alpha_{gt} + \beta \ln \frac{Flow_{jgt}}{Population_{jg(g+r)}} + \varepsilon_{jgt}.$$

The independent variable is the total flow accruing to a cohort relative to the size of the cohort in the state around some modal age r , where this age reflects the typical age of degree completion. The dependent variable is the stock of degree recipients measured years after degree conferral for each cohort relative to the population in the state. It is worth noting that flows here measure the number of individuals in cohort g who eventually receive a degree in state j , regardless of whether the degree recipients were

residents of the state at the time. In the U.S. most, but by no means all, undergraduate students attend college in their home state. Regardless, we are interested in the location of the degree-granting institution, not the residence status of individual receiving the degree. We present estimates for BA and MD degrees.⁶

In addition to the cross-sectional analysis we investigate how changes in cohort-specific flows translate to changes in cohort-specific stocks. We look at changes in the intervals 1960-1970, 1970-1980, and 1980-1990. Here, the focus is on differences in the measures of flows and stocks over ten-year intervals defined for people of the same age referenced by birth cohort g and $g-10$ in a state j . Again, we present the relationship in an elasticity form:

$$(2) \quad \Delta \ln \frac{Stock_{jgt}}{Population_{jgt}} = \alpha_{gt} + \beta \Delta \ln \frac{Flow_{jgt}}{Population_{jg(g+r)}} + \varepsilon_{jgt},$$

where Δ means differences between 1970 and 1960, etc. More specifically, for a variable x_{jgt} , the ten-year difference Δx_{jgt} is defined as

$$\Delta x_{jgt} = x_{j,g,t} - x_{j,g-10,t-10}.$$

This differenced specification captures medium-run dynamic effects rather than long-run differences measured in the cross-section specification. In terms of interpreting estimates as reflecting the causal effects of flows on stocks, these specifications have the advantage of eliminating state-specific fixed effects. Thus, the variation that we hope to consider in identifying our parameters is the extent to which idiosyncratic changes in a state's college degree output have sustained effects on the concentration of college-

⁶ The unit of the static model (1) is the state-cohort cell. As discussed in more detail with the presentation of the empirical results, the inclusion of year effects means that variation across states is what identifies our estimates.

educated workers in the population. Still, one concern is that causality is running in the reverse direction with changes over time in the state-specific demand for college-educated labor feeding back into changes in the fraction of the college-aged population receiving degrees. Thus, just as is true for our cross-sectional estimates, the overtime estimates may exaggerate the causal effect of shifts in flows.

It is natural to assume that the medium-run impact of a shift in flows on stocks will be larger than the long-run impact. After all, the long-run supply (migration) elasticity will be larger the medium-run elasticity. However, when considering the relative magnitude of medium- and long-run equilibrium shifts, it is important to bear in mind that the magnitude of these shifts will depend on both demand and supply parameters. In particular, if capital is relatively immobile (across sectors and across space) in the short run, but mobile in the long run, then it is entirely possible that the short-run impact of labor supply shifts might be small even if the long-run impact were substantial.

Data

The data used in this analysis are from the decennial Census surveys and annual institutional surveys of degrees awarded by colleges and universities conducted by the Department of Education (further details are available in Appendix 1). For BA degrees, we use institutional data on degrees awarded per year and the national distribution of degree recipients by age to calculate the number of degrees awarded to each birth cohort at the state level (these procedures are detailed in the Appendix 1). To obtain measures of per-capita flows for BA degree recipients, we divide these imputed cohort-specific flows

by the population in the state at age 22, calculated from widely available tabulations of the age distribution from the Census Bureau.

The use of national data on the age distribution of BA degree recipients for calculating the BA flow measures at the state-cohort level undoubtedly introduces some measurement error. The bias caused by this error is complicated and reflects a combination of sampling error in the estimated age distribution, which is likely to be classical in form, and the unobserved cross-state variation in the age distribution at degree receipt, which is likely to be non-classical in form. Because the state-specific flows are substantially stable across time, these errors are unlikely to have any significant effect on our cross-sectional estimates. However, this problem could have an appreciable effect on our dynamic estimates. To gauge the magnitude of this problem, we conducted a number of simulations (see Appendix 3), which suggest that the magnitude of the bias introduced by the imputation error is relatively small.

Our ability to organize the MD information by birth cohort mitigates some of the measurement problems associated with the timing of degree receipt for this group. The data for MD degree recipients is from a database maintained by the AMA that records age and other demographic characteristics, institution of degree receipt, and professional employment location. Because we observe this universe in 1980 and 1991, we are able to make comparisons over time as well as across sections.

To estimate the per-capita stock of BA college graduates in a state, we use micro data from the decennial census for years 1960, 1970, 1980, and 1990. We calculate the share of BA recipients in an age group relative to the population size as our stock measure. The 1990 Census provides an advantage over previous decennial files for this

analysis because educational attainment is presented as degree levels rather than years of completed education. For earlier census years (1960-1980), we make the standard assumption that 16 years of completed education equates to the receipt of a baccalaureate-level degree.⁷ For MDs, we use data from the AMA database on degree receipt to measure the numerator and data from the Census to measure the denominator or cohort size.

Section 3: Empirical Stock-Flow Analysis

Concentration of Flows and Stocks

The starting point for the empirical analysis is the consideration of the concentration of flows and stocks across states and the population. We begin by considering those who received degrees between 1966-1985; for BA graduates this reflects the 27-46 age group in 1990, and for MDs the 32-51 age group. The mean flow and stock measures, presented in the first column of Table 1, are indicative of degree receipt, with BA degree recipients nearly 75 times more prevalent than MDs. Across degree types, BA graduates have a much smaller flow dispersion than MDs, as indicated by the lower cross-state coefficient of variation in BA flows.

Figure 2 illustrates the geographic dispersion in degrees awarded in maps of the flow level by states. Colleges in the plains and northeast states are particularly strong

⁷ The 1990 Census identifies both the state in which a person lives and, for those that work, the state in which they work. Earlier Census enumerations either do not identify state of work, or do so for a subset of the sample. For consistency sake, all results we report are based on state of residence. We did, however, replicate our 1990 cross sectional results classifying individuals according to the state in which they work. Switching to state of work made virtually no difference to any of our results.

producers of BAs. States like New Hampshire, Vermont, and Massachusetts have nearly twice the per-capita flow as states like Georgia, South Carolina, and California at the BA level.⁸ There is appreciably more variation across states in the production of degrees of MD degrees. At one extreme, states that are not densely populated such as Montana, Idaho, and Wyoming do not record any institutions awarding MD degrees. At the other extreme, states such as New York, Illinois, and Iowa report relatively high production of MD degrees.

Table 1 also presents the analysis of variance for the stock and flow measures for BA and MD degrees. Decomposing the observed variance for the two decades of state-level observations reveals that the bulk of the variation is consistently across states. For example, at the BA degree level about 77 percent of the observed variation in flows is across states. Such persistence in the difference in the production of BA degrees awarded points to the presence of long-run differences, and in fact, cross-state flow dispersions were similar during the entire 20th century. As seen by comparing the bottom map to the top map in Figure 2, the dispersion of flows of BA degrees in 1929 is remarkably similar to the more recent distribution of flows, and the correlation between the two is 0.5 though the absolute level of dispersion across states is appreciably lower today.

Cross-Sectional Stock – Flow Analysis

We also limit the analysis to the 48 continental states, as data for Alaska and Hawaii are often difficult to obtain in early years and the obvious differences in geographic integration may lead to somewhat different dynamics. In most cases, we present estimates without DC as the unusual political and industrial structure of this area often leaves this case an outlier.

⁸ Recall that these measures reflect the state in which the degrees are obtained, not the state of student residence. Since the significant majority of undergraduates attend college in state, a measure based on the state of residence of students would produce a very similar picture. That said, some of the states that produce a larger number of BAs such as Vermont and Utah are net importers of out-of-state students, while others such as New Jersey, Connecticut, and California, on net, send some students out of state for college.

Although the concentration in the production of university-educated workers is readily evident from measures of dispersion, the analytic question of interest is the impact of flow variation on the stock of college-educated labor across states. At issue in the interpretation of the cross-sectional relationships between flows and stocks is the extent to which the observed variation in the flow measure is exogenous, rather than reflecting essentially permanent differences in labor demand across states.

The strength of the eastern states in the production of BA degrees can be traced to the relatively intensive concentration of private colleges in this part of the country, many formed before the Civil War by denominational organizations. Similarly, the presence of Brigham Young University in Utah accounts for a large fraction of the degrees awarded in this state and it is unlikely that Utah would have the same fraction of college graduates if the Mormons had settled in Wyoming, rather than Utah.

The dispersion of public colleges also reflects the federal political process in the 19th century. Passed by Congress in 1862, the Land Grant College Act (or the first Morrill Act)⁹ provided funds to establish a college in each state. Because many states were rural or sparsely populated, this large-scale federal support for public higher education placed colleges and universities in states that some might have regarded as too small to support a college of efficient size (Jencks and Reisman, 1969). Moreover, the composition and preferences of the population within a state during the early part of the century shaped the willingness of state governments to invest in the expansion of public higher education. Goldin and Katz (1999) suggest that the level of income in a state and the degree of homogeneity (in terms of religion, ethnicity, and income) in the early 20th century were

important indicators of state-supported expansion of colleges and universities. Goldin and Katz also emphasize the importance of demand-side factors: states with large historical concentrations of industries such as mining, manufacturing, and agriculture were most likely to invest in the expansion of public higher education in the early part of the 20th century.

Overall, the distribution and scale of colleges and universities across states reflect a range of factors including the founding of private colleges in the 18th and 19th centuries, the willingness of local populations to support public expenditures on higher education, the introduction of federal support through the land-grant colleges, and the industrial composition of a state. Some of these factors would seem largely exogenous to state labor markets, while others clearly are not. States with industries that have historically hired a disproportionate share of college graduates may be those that have invested in producing a supply to match the demand, thereby contributing to an upward bias in the cross-sectional estimates.

Table 2 presents OLS estimates in elasticity form of the cross-sectional link between flows and stocks, represented by equation (1).¹⁰ Results indicate a modest association between flow and stock for BAs, with an elasticity of 0.3. Plainly, states with a relatively high production of undergraduate students also tend to have relatively high concentrations of the university educated in their working-age populations. Yet, this

⁹ This act granted each state 30,000 acres for each senator and representative in Congress, with the proceeds from this land resource to be used to fund at least one college.

¹⁰ The inclusion of cohort-year dummies implies that the parameter is identified from cross-state variation. Standard errors are estimated allowing for arbitrary heteroskedasticity and state-level clustering in the errors. As can be seen in Table 1, most of the variation in flows is accounted for by permanent cross state differences. As a result, we would have lost very little in terms of precision by estimating a regression of state specific average stocks on average flows. Indeed, such regressions yield parameter and standard error estimates that are virtually identical to the ones we report

relationship is appreciably less than 1:1 and points quite clearly to the importance of migration as a mechanism for mitigating the impact of supply shocks on local economies. Calculating stock-flow relations for different age ranges and at different points in time for all BA degrees (results not shown) underscores the persistence of the basic result.

Comparisons across degree types highlight the quite different labor markets faced by university-educated labor with different types of training. The cross-sectional relationship between the production of MD degrees and the representation of MDs in the population is remarkably weak, with an elasticity estimate very close to zero. Taking the point estimates at face value, an increase of 10 percent in the medical degrees awarded in a state would be expected to produce an increase in the concentration of physicians in the state of less than 1 percent. The weak link between the flow and stock of MDs is consistent with our expectations. Since MDs work primarily in the non-traded sector, the within-state demand for their services will be quite inelastic.

Graphical presentation of flows and stocks helps to sharpen our understanding of these estimates. Each panel represents the stock-flow relationship averaged over the 1966-85 degree cohorts, with the diagonal line distinguishing net importers (above) and net exporters (below). For the stock and flow of BA degrees, states such as California and Connecticut are BA importers while other states like Utah and Vermont consistently export baccalaureate-trained personnel.¹¹ The picture for MDs is striking in the lack of association between flows and stocks.

The analysis of the relationship between stocks and flows among BA recipients

indicates a persistent and significant link between BA degrees awarded and the representation of college-educated people in the state. Nevertheless, all of our estimates suggest that the elasticity of stocks with respect to flows is appreciably less than one and we have argued that is likely that these estimates exaggerate the actual causal effect of flows.

Instrumental variables (IV) estimation provides a strategy to isolate the causal effect of the production of college-educated workers on the long-term stock. For our cross-sectional estimates, where we are considering relatively permanent differences across states, we use historical dimensions of the higher education industry and demographic differences across states to try to isolate factors that affect production today but are exogenous to contemporary developments in the labor market. One potential instrument is the per-capita flow of BA degrees in 1929; a second potential instrument is a measure of historical ethnic diversity, which potentially reflects the willingness to support public expenditure on higher education.¹²

Although IV estimation provides an alternative strategy to isolate causal effects of flows on stocks, many of the historical variables we have considered are also plausibly related to historical differences in industrial structure. If anything, the IV estimates tend to be somewhat larger than the corresponding OLS estimates, though the differences are not statistically significant. While we do not find our IV estimates more credible than our OLS estimates, they do underscore the point that the observed cross-sectional estimates

¹¹ Looking at this picture divided by cohort (not shown) demonstrates some consistency indicative of the measurement of long-run equilibrium, as well as variation over time. For example, the state of Washington shifted from a relative exporter of BA-level workers in the early decades of observation to a relative importer in the 1980s, while the state of Arizona shifted from a relative importer to exporter.

reflect differences persisting throughout the 20th century.

Variation in relative wages across states with the concentration of college-educated workers provides an indicator of the direction of the causal relationship between flows and stocks and the degree of mobility in the labor force. Table 3 presents estimates of the regression of relative wages for college graduates on the concentration of college graduates at the state level for different decennial points of observation. The relative wages were adjusted for demographic characteristics (for more details, see the Appendix 1). The first column uses the observed concentration of college graduates as the explanatory variable. Column (2) uses the aggregate of flows (from 1950 to the indicated year) as an instrument for the stock of college graduates and this flow measure captures variation attributable to differences across states in degree production from higher education.

In an integrated labor market in which labor adjusts fully in location to changes in demand, these coefficients would be uniformly indistinguishable from zero. Yet, particularly in the IV estimates, these estimates are consistently negative, implying an inverse relationship between flows and relative wages. This result is consistent with a situation in which some states have a comparative advantage in producing college graduates while others have a comparative advantage in the use of college-educated labor. Moreover, our estimates suggest that even in the long run, college-educated labor is not

¹² We use a version of the ethnic diversity measure introduced by Alesina, Baqir, and Easterly (1999) in their work on public expenditures. The computation of this index is discussed in the data appendix.

perfectly mobile across states.¹³ College graduates residing in states that produce a relatively large number of college graduates per capita tend to earn relatively little, while college graduates in states that employ a large number of college graduates but do not produce a large number tend to receive something of a wage premium.

The coefficients in the second column of Table 3 can be interpreted as the inverse of the elasticity of relative demand for college-educated labor. The consistency of these estimates relies on the exogeneity of the instruments. If, however, the reported flow measure is endogenous, the reported coefficients will tend to underestimate the causal effect of relative supply on relative wages and, as a result, will tend to overestimate this elasticity of demand.¹⁴ Taking these coefficients at face value (i.e., interpreting them as estimates of the inverse of the elasticity of demand) suggests a within-state relative demand elasticity in the neighborhood of 5. These estimates are all substantially larger than comparable estimates using U.S. times series data (Katz and Murphy, 1992), suggesting considerable reallocation of production across states accounts for cross-state differences in relative stocks. However, it also seems clear that even in the long run, within-state relative demand elasticities are well below infinity. Exogenous, cross-state increases in the supply of college graduates are accommodated by the out-migration of

¹³ There are a variety of plausible reasons for why this might be true. The costs of moving may be sufficiently high for a large enough fraction of the population that the discounted benefits to relocation are not sufficiently high to merit the move. Alternatively, the marginal college graduate may have a preference for living in states that produce a large number of college graduates (this could simply represent a preference for living near friends and relatives). In this case, the earnings differences would represent equalizing differences.

¹⁴ If flows are endogenous, then the regression of stocks on flows will tend to over estimate the causal effect of flows on stocks. Similarly, in this case the regression of relative wages on flows will tend to underestimate the causal effect of flows. The IV estimates are the ratio of these two estimates, and therefore will tend to underestimate the causal effect of stocks on relative wages.

college graduates and the drop in relative wages, as well as by the reallocation of production across sectors.

Dynamic Stock – Flow Analysis

Difference estimates capture changes over a relatively short horizon and thus measure something conceptually different from our cross-section estimates, which reflect permanent cross-state differences in educational capacity. In this regard, we want to know what happens to the stock of college graduates in a state if the degree output of the state's higher education institutions changes at a rate different from the national norm for a short interval.¹⁵ Table 4 presents the means of the decennial log differences in flows and stocks by age and period of observation. As is well known, overall college attendance expanded dramatically into the early 1970s, accounting for the large and positive changes in flows for those in the 25-34 age range during 1960-1970 and 1970-1980. Decreased labor market returns to college education faced by cohorts making educational investments in the mid- and late-1970s contributed to the decline in flows for the 25-34 age group over the interval 1980-1990.

Turning back to Table 1, the analysis of variance numbers indicate that variation within states over time is an appreciably smaller share of the total variance than the cross-sectional differences. However, as Table 4 and Figure 4 indicate, there is still significant cross-state variation in the change in the flows from one decade to the next. Thus, for example, while average per-capita flows for 25-34 year olds increased by roughly 25

¹⁵ The data support this interpretation, as there is not uniformity in the correlation of changes in flows. States that increased relative flows between 1960 and 1970 were not identical to those with relative increases between 1970 and 1980, though there is a positive relationship between the 1970 to 1980 change and the 1980 to 1990 change. Overall, none of these relationships among flows is very strong nor is there evidence that they persist over time.

percent between 1970 and 1980, this growth varied considerably. It was close to 10 percent for states such as Oregon, Utah, Wyoming, and Nebraska; between 36 and 46 percent for Florida, Nevada, Alabama, and Virginia; more than 50 percent for New Mexico; and about 70 percent for Delaware.

In terms of the changes over time, most of the movement is found in the public sector,¹⁶ where there was significant growth both at research universities and among public comprehensive institutions. Because it is the per-capita flow variable we use in our analysis, changes in this variable can reflect movements in either the numerator or the denominator. In fact, our data yield a strong negative correlation between changes over time in the size of the 22-year-old population and changes in per-capita flows. Indeed, regressions of the change in per-capita flows on the change in the size of the cohort suggest that a 10 percent increase in cohort size is associated with 7 percent decrease in per-capita flows. Statistically, cohort size explains about 25 percent of the variation in the change over time in per-capita flows.

The over-time analysis eliminates permanent cross-state differences, including long-run differences in state demand for college-educated workers, which contribute to bias in the estimates of the effects of flows on stocks. Nevertheless, the change over time in per-capita flows could still be endogenous to state-specific changes in the demand for college-educated labor. When thinking about how serious an issue this is, it is important to understand that the variation at issue represents cross-state differences in the growth of flows from one decade to the next. Growth in one decade is typically not followed by growth in the next. Although all states experienced an increase in the fraction of their

¹⁶ Across time within states, there is roughly three times as much variation in per-capita flows in

college-aged population attending and finishing college between 1950 and 1970, the timing and magnitude of these increases varied, suggesting exogenous factors at work. The actions of governors in the sphere of higher education are one such potentially exogenous force.

The expansion of higher education in New York state under the gubernatorial terms of Nelson Rockefeller represents a striking case in point. Few observers early in the Rockefeller administration would have predicted a six-fold increase in state funding for higher education in New York state in the decade between 1956 and 1966, with the increase in New York exceeding the changes in neighboring Connecticut and New Jersey by 60 percent and 45 percent, respectively. Yet, denied a national office with the nomination of Nixon in 1960, Rockefeller threw his considerable personal energy and ambition into capital projects in the state including the transformation of the SUNY system from teachers colleges to a national-level university system.

At the other extreme, Michigan's Governor Milliken reduced state support for higher education by \$50 million in 1983, illustrating that public colleges and universities are politically viable targets for gubernatorial budget slashing when faced with revenue shortfalls (Gove, 1988). Another type of relative contraction in state support is apparent in the tightly constrained growth of southern systems of higher education during the 1960s, as pressure to desegregate higher education may have thinned political support for colleges and universities. In the case of South Carolina, for example, the residuals from a regression predicting flows with state and year effects are negative throughout the 1960s and early 1970s.

public sector than there is in the private sector over the 1966-1985 time period.

While state-level politics have a clear influence on the over-time variation in the outputs of higher education within a state, the strength of this effect varies appreciably across states with the composition of public and private institutions. In states such as California where public institutions constitute the majority provider of higher education, substantial flow accommodations are likely to accompany changes in population. Alternatively, in a state like Massachusetts where higher education has been provided largely by private institutions, accommodations in degree outputs to population growth or political pressure are likely to be more muted. To put this in perspective, in 1998 public institutions awarded 74 percent of the BA degrees earned in California and only 31 percent of those earned in Massachusetts.

These kinds of considerations lead us to infer the presence of considerable exogenous variation in the state-specific changes over time in per-capita flows. Table 5 presents estimates with the decennial change in stock regressed on the decennial change in flow for different age cohorts. These dynamic estimates, reflecting the difference presentation from equation (2), use variations over time within states to identify the effect of flows on stocks. Estimates for relatively recent college graduates – those that are 25-34 years old as of the census years – are shown in the first column. For these cohorts, the difference estimates show significant effects of flows on stocks (0.36 to 0.44) for the 1960-1970 and 1980-1990 intervals, while the estimate for 1970-80 is somewhat weaker. The scatter plots in Figure 4 reveal several outliers. For the 1960-1970 decade, the Dakotas and Nebraska witnessed particularly dramatic increases in degrees conferred relative to the size of the college-age cohorts during the 1950s. For the 1970-80 decade, Delaware and New Mexico are notable for their relative expansion in college graduates

during the 1960s, an increase that yielded relative little growth in the proportion of college-educated workers living in these states. Removing these two states from our calculations produces results for this cohort that are much more in line with results for other cohorts and that show statistically and quantitatively large associations between changes in flows and changes in stocks [0.28 (0.08)].¹⁷ However, outliers usually contain valuable information. In Delaware and New Mexico, despite enormous increases in the per-capita share of college graduates, the stocks did not increase some years down the line. These outliers seem to confirm our tabulations, showing that flows have at best a moderate effect stocks.¹⁸

Table 5 also presents results for older age groups that would typically have graduated from college more than 10 years prior to the year in which we observe them (columns 2 and 3). These results indicate that the relationship between flows and stocks tends to diminish somewhat as cohorts age, with the elasticity declining to about 0.22 (0.08) for those in the 35-44 age group and to .08 (0.07) for those in the 45-52 age group. When thinking about this diaspora of college graduates, it is important to bear in mind that the growth in flows in one decade is not generally “ratified” by increases in flows in following decades. Thus, the impact of a change in flows on stocks two to three decades later is conceptually distinct from the long-run impact of a change in flows (i.e., the kind of quantity we were attempting to estimate using the cross-state variation in flows).

¹⁷ For those in the 25-34 age cohort, difference estimates for other cohorts include 0.30(0.09) for 1970-1960 and 0.41(0.06) for 1990-1980 for regressions limited to 46 states and excluding DC, Delaware, and New Mexico.

¹⁸ Here, and in other places, we see evidence that the impact of flows on stocks in states that are small either in terms of land area or population, tends to be particularly weak. We tried testing such hypotheses statistically by including interaction terms in our models. Generally speaking, the estimates on the interaction terms suggested that the smaller a state the weaker is the association between flows and stocks. However, the estimated interaction terms were generally not statistically significant. Given the

The graphs in Figure 4 illustrate the dynamics underlying the regression results. States above the 45-degree line are cases in which the change in the stocks of college-educated workers exceeded the change in flows of college graduates, which means these states increased net imports of college graduates. In turn, states below the line tended to shift to exporting college graduates. The cross-state patterns vary considerably over time, with the shifts observed in the 1970-1960 period markedly larger than those observed for 1990-1980.

To address the concern that these first difference estimates do not necessarily reflect the causal effect of flow shifts, the optimal fix would be to use exogenous factors that have changed over time as instruments for changes in flows in our difference specification. Tuition rates and the level of state appropriations for higher education both have some effect on college completion rates. However, the effects are not strong enough to provide us with IV estimates of any reasonable degree of precision.

As an alternative, we have tried controlling for various factors that might confound the effect of changes in flows on changes in stocks. The bottom two panels of Table 5 present estimates that include state-level measures of population and employment as explanatory variables. In addition to affecting the proportion of any cohort graduating from college, cohort size may also affect the location decisions of these individuals. There is ample evidence indicates that large cohorts tend to face worse labor market prospects than do small cohorts, with this being particularly true for college graduates (Freeman, 1979; Welch, 1979). As a result, college graduates in large (state-specific) cohorts will tend to have an incentive to move out of state.

sample size we are dealing with (effectively 48 observations), this was hardly surprising.

Including cohort size at the time of college graduation (the population aged 22, measured in logs) as an additional covariate provides a direct control for the effect of cohort size on the fraction of college graduates in the state. These results are shown in the middle panel of Table 5 and, while the additional covariate places downward pressure on the coefficient, this effect is relatively small. Including direct measures of demand captured by the employment level in the reduced-form differenced regression (bottom panel of Table 5) addresses the concern that the estimated elasticity reflects the effect of local demand shocks on flows. Point estimates in this specification change only slightly from the original specification. Overall, these estimates are very similar to the original difference estimates.

Turning to dynamic estimates in the MD-degree market, the evidence presented in Table 6 indicates a positive and significant relationship between changes in flows and changes in stocks. Scatter plots of the data presented in Figure 5 reveal a clear and compelling story. States that had the largest changes in flows tended to be states like West Virginia and South Dakota that may have been underserved in medical care at the beginning of the interval. As such, the addition of medical schools to two West Virginia universities could have been a policy remedy to increase the supply of doctors in the state.¹⁹ Both institutions have mission statements that address the need to provide physicians and medical personnel for underserved areas and make explicit reference to recruiting students from rural West Virginia and placing graduates in clinical practices to improve health care in West Virginia. In the context of our model, it is likely that the

¹⁹ These were the West Virginia University School of Medicine (part of the Robert C. Byrd Health Sciences Center) and the Joan C. Edwards School of Medicine at Marshall University. West Virginia University awarded its first MD in 1962 and Marshall University established its medical school in 1977.

medium-term effects of changing the production of MDs within a state may be relatively large as the additional MDs produced in a state like West Virginia include many people who are from West Virginia and have a preference for remaining in the state. Still the absolute magnitudes of the coefficients are small (0.2) and indicate that for each ten additional physicians trained in the state, only about 2 will remain in the state's population.

Section 4: Conclusion

The empirical evidence in this analysis points to a modest relationship between degree production in the education market and the concentration of college-educated workers in a state's population in the long run. For BA degrees, we estimate the long-term elasticity between stocks and flows to be no greater than 0.3. In contrast, for MDs we find little relationship between where Physicians are trained and where these practice. The contrast between our results for BAs and our results for MDs suggests that the nature of demand in the labor market is a substantial determinant of the stock-flow relationship, with the relatively inelastic demand for MDs within states in long-term equilibrium contributes to the wide dispersion of MD graduates across states and the relatively weak link between flows and stocks. For MDs, dynamic estimates are somewhat larger than long-run estimates (though still smaller than the corresponding estimates for all BAs) and analysis of specific cases points to the influence of public policies intended to expand the supply of MDs in underserved areas.

Our results indicate that state policymakers have only a modest capacity to influence the human capital levels of their populations by investing in higher education

degree outputs. Although, presumably, states have a number of reasons to invest in higher education, the mobility of college-educated labor reduces their incentive to make such investments (Justman and Thisse, 1997). Indeed, some evidence suggests that the degree of mobility affects states' investment behavior. States with a higher out-migration of college graduates tend to have lower per-capita education expenditures (Clotfelter, 1976; Strathman, 1994). What is more, the magnitude of the subsidies states provide to medical students appears to be quite low relative to the subsidies provided to undergraduate education.²⁰

Our estimates are also suggestive of how state economies adjust to supply shocks in the U.S. The labor economics literature (e.g., Blanchard and Katz, 1992; Borjas, Freeman, and Katz 1997) has argued for the importance of migration as a means that states have of adjusting to macroeconomic shocks. In contrast, trade economists (e.g., Davis, Weinstein, Bradford, and Shimpo, 1997; Hanson and Slaughter, forthcoming) have emphasized the important of output mix shifts. Our results suggest that, at least in the U.S., migration does have a central role in mitigating the effect of labor supply shifts, but clearly other adjustment processes are also at work. For BA degree recipients, who are relatively likely to work in the traded goods sector, the within-state relative demand for college-educated labor seems to be substantially higher than the one for the country as

²⁰ Given the joint nature of production, it is difficult to estimate directly state subsidies to either undergraduate or graduate education. The difference between private and public institutions in the tuition charged to medical students is much smaller than the difference private and public institutions in the tuition charged to undergraduate students. For universities offering both BA and MD degrees, private universities charge on average a tuition of \$25,655 in MD programs and \$18,235 in undergraduate programs. At public universities, the average tuition charged to medical students is \$10,398 for residents and \$22,353 for non-residents, while the average tuition charged to undergraduates is \$3,140 for residents and \$8,668 for non-residents. (These data are from academic year 1996-97. Data for tuition and fees in MD programs are from the Association of American Medical Colleges and the corresponding undergraduate data are from the IPEDS survey.)

a whole, suggesting that supply shifts induce the reallocation of production across sectors. What is more, even in the long run, wage differentials between college and high school graduates seem to be relatively small in states that produce a large number of BAs on a per-capita basis, suggesting the college-educated labor is less than perfectly mobile across states.

References

- Alesina, Alberto, Reza Baqir and William Easterly. 1999. "Public Goods and Ethnic Divisions." *The Quarterly Journal of Economics*. Volume 114, No. 4.
- Acemoglu, Daron and Joshua Angrist. 2000. "How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws," *NBER Macroeconomics Annual*, pp. 9-59
- Bartik, Timothy. 1991. *Who Benefits from State and Local Economic Development Policies?* W.E. UpJohn Institute, Kalamazoo, MI.
- Blanchard, Olivier and Lawrence Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity*, No. 1, pp 1-61.
- Borjas, George J.; Richard Freeman, and Lawrence Katz. 1997. "How Much Do Immigration and Trade Affect Labor Market Outcomes?" *Brookings Papers on Economic Activity* v0 n1 1-67.
- Bound, John and Harry Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics*, 18:20-54.
- Clotfelter, Charles. 1976. "Public Spending for Higher Education: An Empirical Test of Two Hypotheses." *Public Finance*, 31, 174-194.
- Davis, Donald; David E. Weinstein, Scott C. Bradford, and Kazushige Shimpo. 1997. "Using International and Japanese Regional Data to Determine When the Factor Abundance Theory of Trade Works." *The American Economic Review*, Vol. 87, No. 3. (June), pp. 421-446.
- Freeman, Richard. 1979. "The Effect of Demographic Factors on the Age-Earnings Profile in the U.S." *Journal of Human Resources* XIV:3.
- Glaeser, Edward L., Jose A. Scheinkman, and Andrei Shleifer. 1995. "Economic Growth in a Cross Section of Cities," *Journal of Monetary Economics*, 36:117-143.
- Glendon, Spencer, 1998. "Urban Life Cycles." Harvard University Mimeo.
- Goldin, Claudia and Larry Katz. 1999. "The Shaping of Higher Education: The Formative Years in the United States, 1890 to 1940." *The Journal of Economic Perspectives*, Vol. 13, No. 1 (Winter).
- Gove, Samuel. 1988. "Governors and Higher Education: A Stock-Taking" in *Governors and Higher Education*. Gove and Beyle (ed.) Denver: Education Commission of the States.

Hamermesh, Daniel. 1986. "The Demand for Labor in the Long Run" in O. Ashenfelter and R. Layard, eds., *Handbook of Labor Economics*, North-Holland Press.

Hanson, Gordon and Matthew Slaughter. forthcoming. "Labor Market Adjustments in Open Economies: Evidence from US States." *Journal of International Economics*.

Jencks, Christopher and David Reisman. 1969. *The Academic Revolution*. Anchor Books: Garden City, N.Y.

Jones, Ronald. 1965. "The Structure of General Equilibrium Models." *Journal of Political Economy* 73 (December): 225-36.

Justman, Moshe and Jacques-Francois Thisse. 1997. "Implications of the mobility of skilled labor for public funding of higher education." *Economic Letters* 55, pp. 409-412.

Katz, Lawrence and Kevin Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics*, Vol. 107, No. 1. pp. 35-78.

Long, Larry E. 1988. *Migration and residential mobility in the United States*. New York: Russell Sage Foundation.

Moretti, Enrico. 1999. "Estimating the Social Return to Education: Evidence from Repeated Cross-sectional and Longitudinal Data." Center for Labor Economics, Working Paper 21.

Newhouse, Joseph P., et al. 1982. "Does the Geographical Distribution of Physicians Reflect Market Failure?" *Bell Journal of Economics* v13, n2 (Autumn 1982): 493-505.

Rauch, James E. 1993. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities." *Journal of Urban Economics* v34 n3 380-400.

Strathman, James. 1994. "Migration, Benefit Spillovers and State Support of Higher Education." *Urban Studies*. Vol. 31, 6 (June).

Welch, Finis. 1979. "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust." *The Journal of Political Economy*, Vol. 87, No. 5, Part 2: Education and Income Distribution. pp. S65-S97.

Working, E.J. 1927. "What Do Statistical 'Demand Curves' Show?" *Quarterly Journal of Economics*, Vol. 41, No. 2. pp. 212-235.

Table 1: Stock and Flow Summary Statistics Cohorts, 1966 –1985 degree cohorts

		# per 1000	Cross-State CV	Analysis of Variance		
				State	Cohort	Within
BA	Flow	256.09	1.1	0.77	0.12	0.11
	Stock	243.29	0.8	0.69	0.21	0.10
MD	Flow	3.47	2.7	0.87	0.05	0.08
	Stock	4.33	0.8	0.53	0.26	0.21

Notes: “Flow” data represent the number of degree recipients from a state divided by the age-appropriate population. Measures of the number of BA degrees awarded at the cohort-state level were estimated from year-state level institutional data and the national yearly age distribution of new college graduates. Cohort-state level numbers of MD degrees come from the AMA database. “Stock” data are the number of college graduates living the state in 1990 in the appropriate age group, relative to total population of the age group in the state. See Appendix 1 for additional information on the data used in this analysis. The CV, or coefficient of variation, is the partial coefficient of variation reflecting between state variation in flows and stocks.

Table 2: Cross-section stock-flow elasticities, 1966-85 Degree Cohorts

	Elasticity		Elasticity
	Ln-Ln (1)	Linear (2)	from (2) (3)
BA	0.34 (0.08)	0.30 (0.07)	0.32 (0.08)
MD		0.08 (0.04)	0.07 (0.04)

Notes: The regressions include data for 48 continental states and exclude the District of Columbia. The stock and flow variables correspond to the degree type listed in the first column. The regressions include cohort-specific fixed effects and correspond to the specification in equation (1) in the text. Standard errors are calculated using the method of Huber-White and allow for arbitrary clustering at the state level. Flow and stock variables are associated with the 1990 Census year.

Table 3: Estimates of effect on relative wages of concentrations of college-educated labor

	OLS	IV Aggregate Flow 1950-t
	(1)	(2)
1960	-0.02 (0.06)	-0.38 (0.31)
1970	-0.06 (0.03)	-0.28 (0.15)
1980	-0.04 (0.02)	-0.16 (0.08)
1990	-0.03 (0.02)	-0.19 (0.09)

Notes: Data used in these calculations are from the indicated Decennial Census files. The dependent variable in the cross-section measure is the regression-adjusted state-specific measure of the difference in the log of college and high school wages. The independent variable is the log of the ratio of college degree recipients to those with a high school degree or less.

Table 4: Means of difference measures of flow and stock

Census Years	Age Range	Ln Difference of BA Flow		Ln Difference of BA Stock	
		Mean	Std Dev	Mean	Std Dev
1970-60	25-34	0.41	0.21	0.37	0.11
1980-70	25-34	0.28	0.18	0.45	0.09
1990-80	25-34	-0.10	0.16	-0.08	0.10
1980-70	35-44	0.46	0.19	0.41	0.09
1990-80	35-44	0.27	0.18	0.36	0.08
1990-80	45-52	0.46	0.19	0.36	0.08

Notes: Table entries reflect the log difference over the indicated decade of averages of flows and stocks for the indicated age ranges (where the age corresponds to the base year).

Table 5: Dynamic estimates of the effects of flows on stocks

OLS, 48 States			
	25-34	35-44	45-52
	(1)	(2)	(3)
1960-70	0.36 (0.09)		
1970-80	0.13 (0.10)	0.22 (0.08)	
1980-90	0.44 (0.06)	0.22 (0.08)	0.08 (0.07)
48 States, Population			
	25-34	35-44	45-52
1960-70	0.31 (0.12)		
1970-80	0.13 (0.11)	0.17 (0.09)	
1980-90	0.40 (0.05)	0.22 (0.08)	0.06 (0.08)
48 States, Population, Empl.			
	25-34	35-44	45-52
1960-70	0.33 (0.10)		
1970-80	0.14 (0.11)	0.20 (0.08)	
1980-90	0.42 (0.05)	0.23 (0.08)	0.07 (0.08)

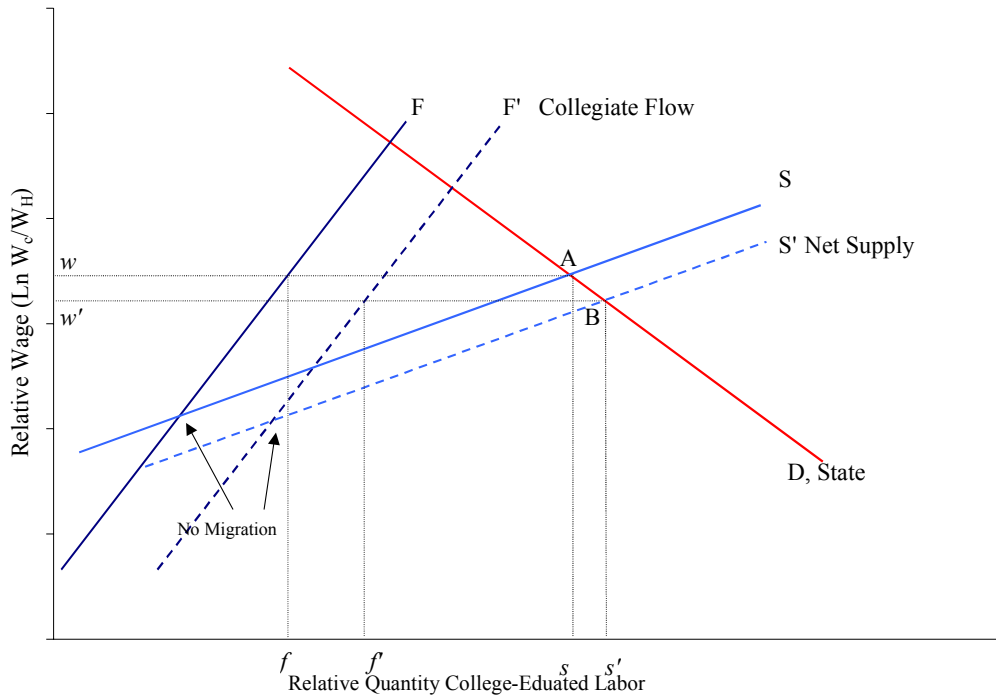
Notes: The regressions include data from the 48 continental states. The regressions are specified with differenced observations for each year within the decennial interval with the inclusion of birth cohort fixed effects and correspond to the specification presented in equation (2). Standard errors are calculated using the method of Huber-White and allow for arbitrary clustering at the state level.

Table 6: Dynamic estimates for MDs, 1991-1980

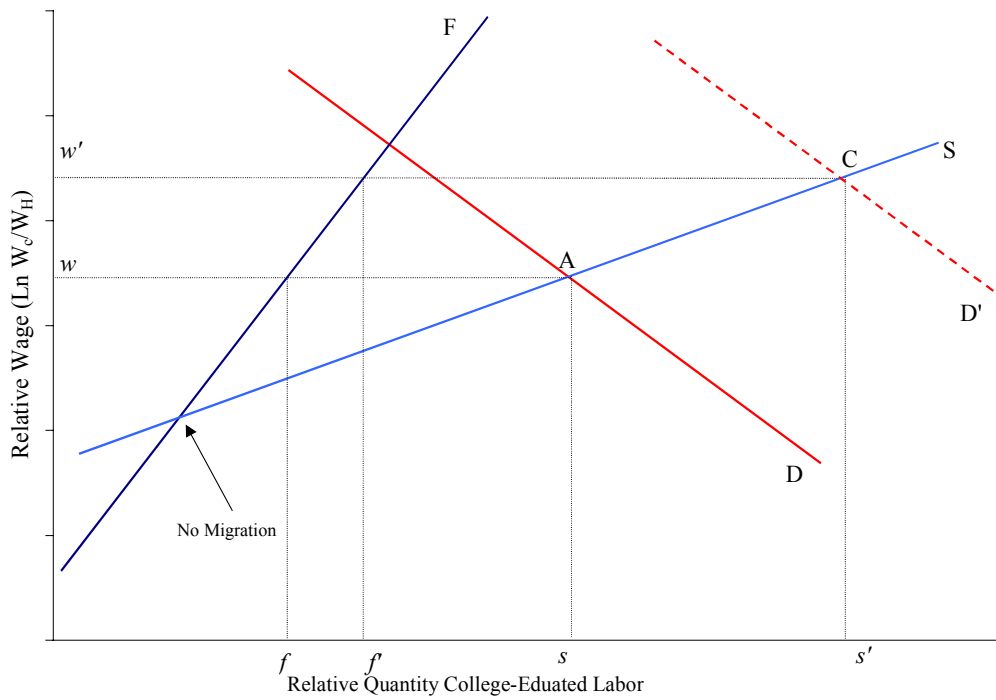
Field and Cohort	Elasticity	
	Linear	from (1)
	(1)	(2)
<i>Medical Doctors</i>		
35-44	0.22 (0.06)	0.17 (0.04)
45-54	0.25 (0.07)	0.16 (0.05)

Notes: The regressions include data from the 48 continental states. The regressions are specified with differenced observations for each year within the decennial interval with the inclusion of year fixed effects and correspond to the specification presented in equation (2). Residents are not included in the tabulations. Standard errors are calculated using the method of Huber-White and allow for arbitrary clustering at the state level.

Figure 1: State-level adjustments to changes in flows.

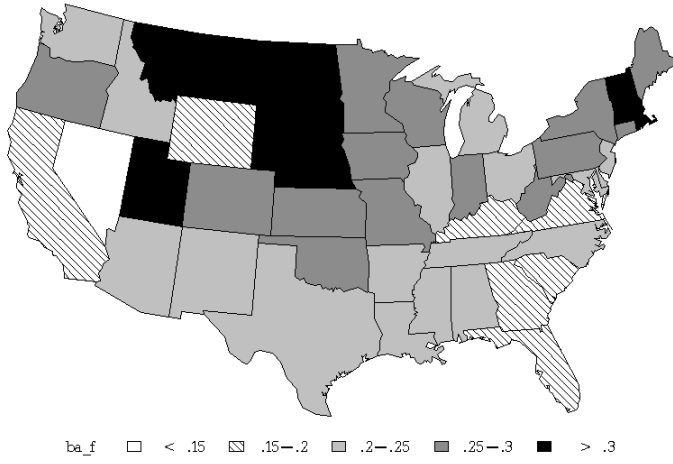


Panel 1

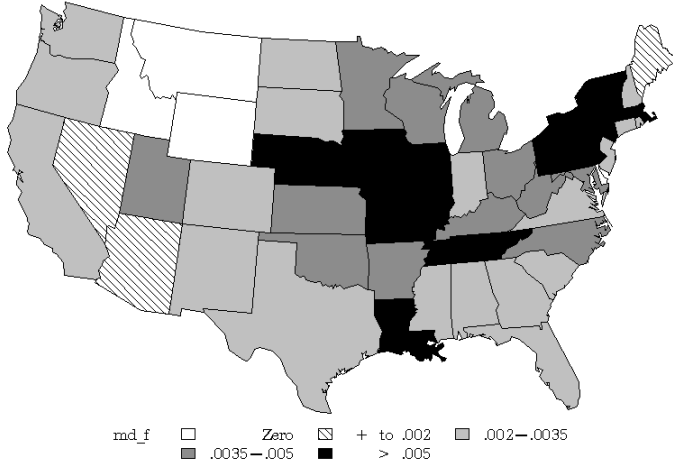


Panel 2

Figure 2: Maps of geographic dispersion of flow variables
 Ratio of BA Flow to Population



Ratio of MD Flow to Population



BA flow per person age 22, 1929

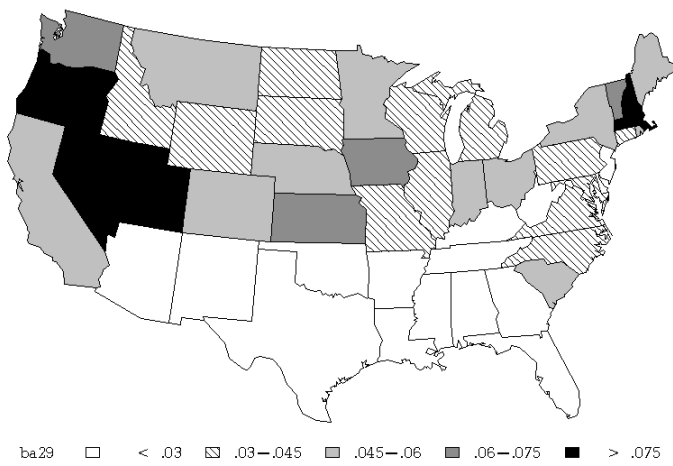


Figure 3: Stocks and flows of degrees awarded relative to cohort size, 1966-85 degree cohorts

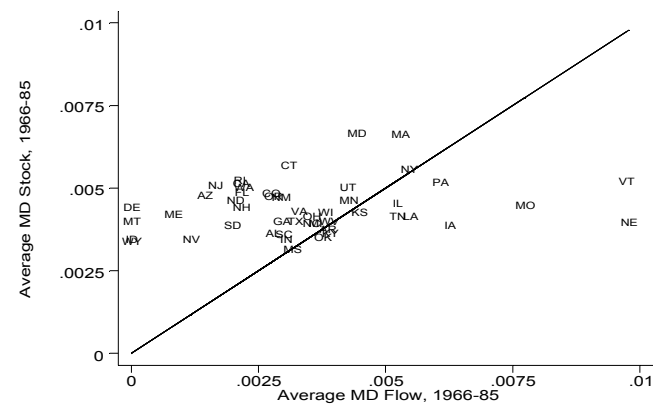
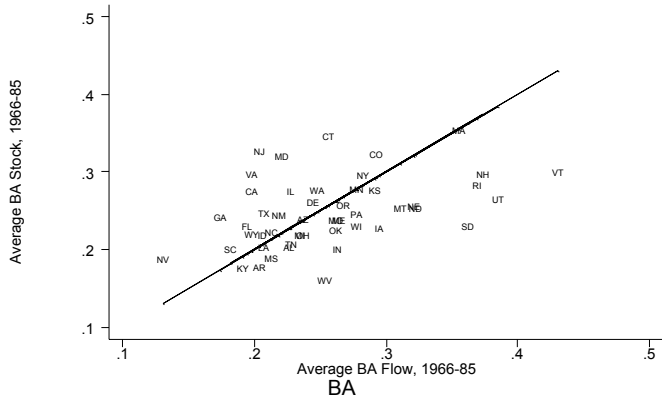


Figure 4: Changes in flows and stocks relative to cohort size, BA age group and period

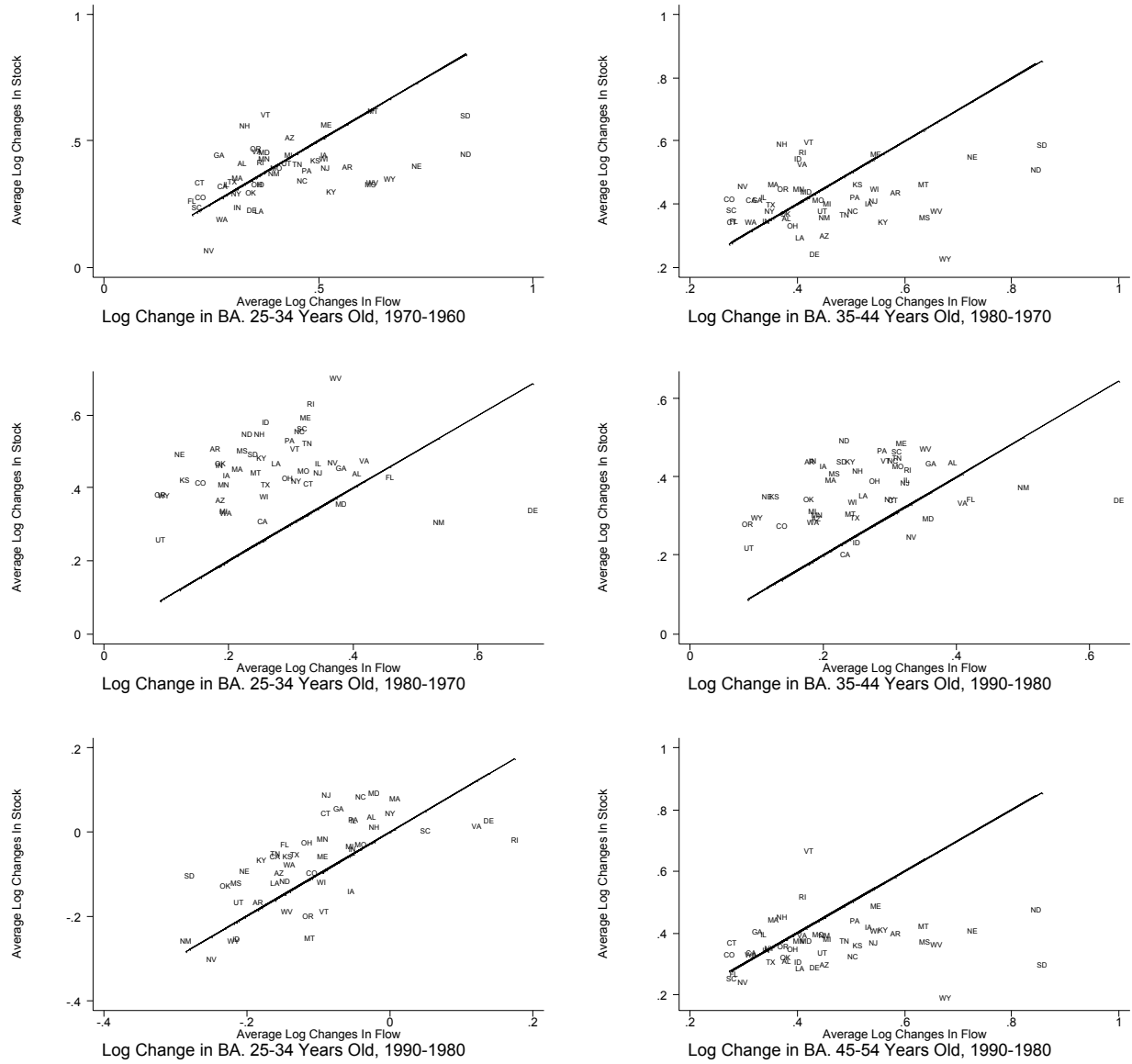
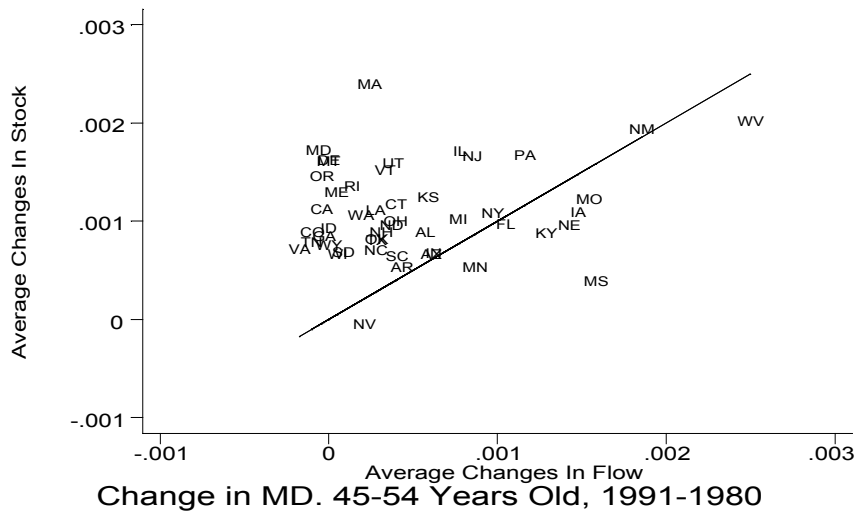
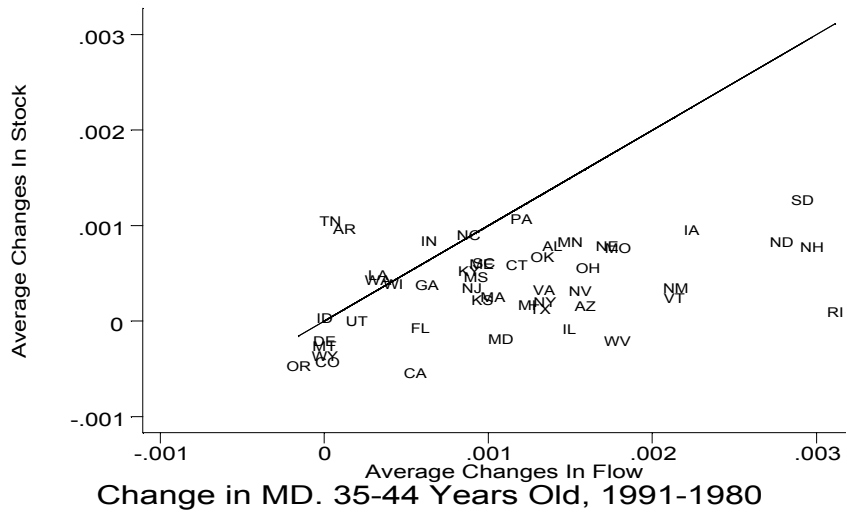


Figure 5: Changes in flows and stocks relative to cohort size, MD degrees



Appendix 1. Data Appendix

The primary sources of data for this analysis are: the decennial Census files, the decennial Census publications, population estimates by the Census Bureau, the October CPS files, institutional surveys of degrees awarded, and the AMA Physician Professional Data. Appendix Table A1.1 lists specific references.

In the paper, we analyze two types of degrees: BAs and MDs. The first part of the appendix is organized by degree type. Steps that apply to both types are discussed with BA degrees.

BA Degrees

Flows

We wish to compute the *per capita flow* of college graduates for each state and birth cohort, and the *per capita stock* of college graduates for each state and cohort for the census years 1960, 1970, 1980, and 1990.

A major measurement problem arises from the fact that, in our baseline data, the *stock* variable is defined for birth cohorts, while the *flow* refers to the year of graduation. Therefore, we had to estimate *flows* for birth cohorts from the degree-year data. We do not observe the flow of BA degrees awarded to members of birth cohorts directly, but only the sum of degrees awarded to cohorts of different ages. While age 22 is the modal age group for BA recipients, not all BA degree recipients are this age.

First, in each year, we estimated the number of college degrees conferred for different ages. Using micro data from the October Current Population Survey, we estimated the age distribution of college seniors and then estimated the number of college degrees conferred to individuals in each birth cohort each year. Micro data from the October CPS are available from 1969 to the present. For years prior to 1969 we assumed the age distribution of college seniors was similar to the age distribution in 1969. Since the age distribution corresponds to the age distribution in October, while most individuals receive their degrees in June, we estimated the spring age distribution by assuming that half of each cohort was a year younger at that time (basically assuming uniform distribution of month of birth).

To estimate the number of degrees conferred by cohort in each state, we multiplied the age distribution of the BA degrees conferred by the annual number of degrees awarded in each state as reported in institutional surveys of colleges and universities. We then assigned the different year-by-age numbers to the corresponding cohorts (defined by year of birth) in each state. Thus, for example, if we estimated that 40 percent of graduating seniors in 1980 were 22, we would assign 40 percent of the degrees conferred in each state in 1980 to that cohort.²¹

²¹ Since we did not have degrees conferred data for years prior to 1950, our per capita flow estimates are truncated for those cohorts that turned 22 in the early 1950s. The (extrapolated) age

Third, we divided this absolute number by the size of the cohort in the state. Since the modal year of college completion is 22, we used estimates of the size of the cohort at age 22 for this purpose. The U.S. Census Bureau reports population estimates by state and single year of age for 1970 and later. For the years prior to 1970, we estimated the share of 22 years old in the total population in a state by a weighted average of the corresponding birth cohorts from the two closest census figures, the weights being inversely proportional to the distance from the given census year.²² We then used these shares and the total population of the states in the corresponding years to estimate the number of 22 year olds.

The per capita flow data is the ratio of these two estimates: the estimated number of degrees conferred for a given birth-cohort in a given year, divided by the size of that cohort in the state when they were 22 years old.

In the analysis, we use stock variables referring to different years: 1960, 1970, 1980, and 1990. When analyzing flows with these different stocks, one does not want to include in the flows those college graduates who received their degrees after the year of the stock. Therefore, we carried out the estimation of the flows four times, each corresponding to one stock variable.

Stocks

To estimate the per-capita stock of college graduates in a state we used micro data from the decennial census for the years 1960, 1970, 1980, and 1990. For 1960, the largest sample available represents a 1 percent random sample of the population. For 1970, the largest sample for which state of residence is available represents a 2 percent random sample of the population. For 1980 and 1990, the samples we used represent 5 percent of the population. In 1990, the census asked about the highest degree received by an individual. We assume that all those who identified themselves as having a Bachelors, Masters, Professional, or Doctorate degree were college graduates. For earlier years, the census asked how many years of college a person had completed. For these years, we assumed that anyone who completed 4 or more years of college was a college graduate.

The 1980 and 1990 censuses allow one to identify state of work as well as state of residence. We did separate tabulations using data organized by state of residence, but found it made very little difference whether we identified individuals by their state of residence or their state of work.

MD Degrees

distribution of the graduates suggests that approximately 70 per cent of a cohort graduated at the age of 22 or before. This means that 30 percent of those who turned 22 in 1950 received their degree before 1950, for the time-period we have no data. For similar reason, we don't have degrees conferred data for 20 per cent of the cohort that turned 22 in 1951, and for 15 per cent of those that turned 22 in 1952, and so on.

²² For example, the share of 22 years old in the population of a given state in 1963 was estimated by 0.3 times the share of the 29 years old in the 1970 census plus 0.7 times the share of the 19 years old in the 1960 census.

For MD degree recipients, we constructed degree flows from the AMA Physician Professional Data. This is a comprehensive source of information on U.S. physicians, including both members and non-members of the AMA. The file includes information on date of birth and medical school for each physician. For physicians trained at U.S. medical schools, we used the year-end files for 1980 and 1991 to construct an estimate of degree flows by state and year of birth. We matched these degree flows – by State and year of birth – with population in the five-year age group 25-29. (About 80 percent of the physicians in the 1991 AMA file received their MD between the ages of 25 and 29.)

We calculated the population age 25-29 in each state and year using Census Bureau data and followed the imputation strategy discussed above under BA degrees in the years prior to 1970 when single year age tabulations were unavailable.

To estimate the per-capita stock of MD's in 1980 and 1991, we tabulated the number of physicians by State and year of birth. For our stock measure, we dropped those who are not actively practicing medicine and those in residencies. To put these stocks in per-capita terms, we divide them by the population in 1980 or 1991, by State and year of birth.

Other Variables

Birth Cohort Size: Data on the size of a birth cohort for each state from 1928 – 1970 (these are the cohorts that would have been 22 between 1950 and 1992) were entered from vital statistics data distributed by the National Center for Health Statistics. The original data came from birth registrations.

BAs in 1929: Counts are from Table 4a “Summary of degrees conferred in 1929-30” of the *Biennial Survey of Education, 1928-1930*.

Racial Disparity Index: Calculated to replicate the variable used by Alesina, Baqir, and Easterly (1999). In particular, we computed the index for a state as: $Q_{kt} = 1 - \sum_k s_{jkt}^2$, where s_{jkt} is the share of the k th racial group in the j th state in year t . Following Alesina, Baqir, and Easterly (1999), we categorized individuals as White, Black, American Indian, Asian, or other. While we computed the index for all census years between 1960 and 1990, though the results mentioned in the text were based on the 1960 values of the index. With five groups, the index has a possible range from 0.0 to 0.8. In fact, in our data the index ranges from 0.003 for Vermont to 0.49 for Mississippi, with an average value of 0.17.

Relative wages: The adjusted average relative wage measures are computed as the return to exactly a BA Degree (or 16 years of completed education) from state-specific hourly wage regressions with a full set of controls for demographic and labor force characteristics including, race, sex, part-time status, and potential experience specified as a quartic.

Table A1.1: Data sources

Source	Census of Population and Housing, United States Department of Commerce. Bureau of the Census.	Use
<p>1960 Public-Use Sample: 1 Percent Sample 1970 Public-Use Sample: 1 Percent Sample, 15 Percent State Questionnaire 1970 Public-Use Sample: 1 Percent Sample, 5 Percent State Questionnaire 1980 Public-Use Sample: 5 Percent Sample 1990 Public-Use Sample: 5 Percent Sample</p>	<p>Bureau of the Census.</p>	<p>Census data provide stock measures of educational attainment by state. Occupational codes such as teacher and physician permit the identification of skills beyond years of educational attainment.</p>
<p>U.S. Census Bureau 1950, 1960, and 1970 State Volumes</p>	<p>U.S. Census Bureau Population Estimates (http://www.census.gov/population/www/estimates/st_stts.html)</p>	<p>Population estimates by state and single year of age, in 1950, 1960, and 1970. Used for estimating cohort size per state at age 22 and age 25-29, for years 1950-1970.</p>
<p>Current Population Survey, October, 1969-1990. U.S. Census Bureau and Bureau of Labor Statistics</p>	<p>U.S. Census Bureau Population Estimates (http://www.census.gov/population/www/estimates/popest.html)</p>	<p>Used to estimate the age distribution of college seniors, 1969-1990.</p>
<p><i>Higher Education General Information Survey (HEGIS)</i>: Earned Degrees, 1967-1968 through 1985-1986. U.S. Department of Education, National Center for Education Statistics.</p>	<p>U.S. Department of Education, National Center for Education Statistics.</p>	<p>Data on degrees conferred at the levels of Associate, BA, and First Professional collected annually through institutional surveys conducted by the Department of Education. Data contain detailed information on field of study.</p>
<p><i>Integrated Postsecondary Education Data System (IPEDS)</i>: Earned Degrees, 1986-1987 to present. U.S. Department of Education, National Center for Education Statistics.</p>	<p>U.S. Department of Education, National Center for Education Statistics.</p>	<p>Data contain detailed information on field of study.</p>
<p>“Earned Degrees Conferred,” Department of Education annual, 1950-1966.</p>	<p>Department of Education annual, 1950-1966.</p>	<p>Department of Education annual, 1950-1966.</p>
<p>Physician Professional Data, 1980 and 1991 year-end files, American Medical Association.</p>	<p>1980 and 1991 year-end files, American Medical Association.</p>	<p>Data contain information on date of birth, medical school, and current state of residence for physicians practicing in the U.S. We use these data to construct stocks and flows by State and birth cohort.</p>

Appendix 2. Derivation of Theoretical Model

To aid in the interpretation of our estimates we develop a conceptual model of the demand and supply in the market for college-educated workers. The purpose of the model is twofold. First, the model places the degree outcomes of higher education at the state level in the context of supply and demand in the labor market, thereby guiding the economic interpretation of our estimates. The model also helps us to understand the nature and direction of the bias that arises when we use the cross-sectional association between stocks and flows to make inferences about the causal effect of flows on stocks.

In the text, we present a graphical version of the model and we develop it algebraically in this appendix. The model we present captures changes within a state in a partial equilibrium context, implicitly treating the state as a small economy. In this regard, wages outside the state are assumed to be exogenous and not affected by migration. We have confirmed that the qualitative implications of the model are robust to the partial equilibrium assumptions using a parameterized general equilibrium model of two equally large states. After setting forth a general framework for labor supply and labor demand, we solve for equilibrium under a set of simplifying assumptions. We then introduce changes in the equilibrium generated by shifts in flows of college graduates and examine the implications for the measurement of the effects of flows on stocks. In the final part of the analysis, we relax these simplifying assumptions and demonstrate that the basic inferences are robust to such simplifications.

Labor Demand

The demand side of the model is based on Jones (1965) and is very similar to Bound and Holzer (2000). Using constant returns to scale technologies, firms produce both for export, X , and for local consumption, Q , using college (L_c) and high school (L_h) educated labor. The area also imports M . X , Q and M should be thought of as composite commodities. The sectors that produce these goods and services each consist of many firms in many industries. Labor is perfectly mobile both between sectors and between firms within sector. We assume that residents of the state have no effective demand for X and produce no close substitutes for M . The first of these assumptions is not essential, but simplifies the model considerably. We assume that factor and product markets are competitive.

Let γ_{ij} represent the share of the i th skill group's workforce employed in the j th sector ($\sum_j \gamma_{ij} = 1$), and S_{ij} the i th skill group's share of the total wage bill in the j th sector ($\sum_i S_{ij} = 1$). It will be useful to have notation for the differences in factor intensities. Define $d\gamma = |\gamma_{hx} - \gamma_{cx}| = |\gamma_{cq} - \gamma_{hq}|$ and $dS = |S_{hx} - S_{hq}| = |S_{cq} - S_{cx}|$.

If we let W_c and W_h represent the wages of skilled and unskilled labor, p_x and p_q the prices of exports and local consumption goods (in terms of the price of imports), and dots over variables the percent changes in those variables, then:

$$(A2.1) \quad (\gamma_{cq}\dot{Q} + \gamma_{cx}\dot{X}) - \dot{L}_c = \delta_c (\dot{W}_c - \dot{W}_h)$$

$$(A2.2) \quad (\gamma_{hq}\dot{Q} + \gamma_{hx}\dot{X}) - \dot{L}_h = -\delta_h(\dot{W}_c - \dot{W}_h)$$

where δ_i represents the aggregate percentage saving in the i th input at unchanged outputs associated with a 1% rise in the relative wages of the i th input. The δ 's are themselves functions of the elasticities of substitution between the two labor inputs in the two sectors: $\delta_h = \gamma_{hx}S_{cx}\sigma_x + \gamma_{hq}S_{cq}\sigma_q$ and $\delta_c = \gamma_{cx}S_{hx}\sigma_x + \gamma_{cq}S_{hq}\sigma_q$, where σ_x and σ_q represent the elasticity of substitution between skilled and unskilled labor in the export and non-traded goods sectors, respectively. When relative wages change, substitution occurs both within and between industries within a sector, and σ_x and σ_q reflect both types of substitution. We expect both σ_x and σ_q to be above 1 (Hamermesh, 1986). Also, since shifts between firms (industries) in the export sector are likely to be substantially larger than in the local goods sector, we expect $\sigma_x \gg \sigma_q$.

The terms in parentheses on the left of (A2.1) and (A2.2) represent shifts in the demand for the two labor inputs. Subtracting (A2.2) from (A2.1) gives:

$$(A2.3) \quad (d\gamma\dot{Q} - d\gamma\dot{X}) - (\dot{L}_c - \dot{L}_h) = (\delta_c + \delta_h)(\dot{W}_c - \dot{W}_h).$$

Shifts in the industrial composition of a local area or shifts in the composition of the workforce affect the relative wages of skilled and unskilled workers in similar ways.

Output prices are a share-weighted linear combination of the shift in wages of skilled and unskilled labor inputs:

$$(A2.4) \quad S_{cq}\dot{W}_c + S_{hq}\dot{W}_h = \dot{p}_q$$

$$(A2.5) \quad S_{cx}\dot{W}_c + S_{hx}\dot{W}_h = \dot{p}_x.$$

Equations (A2.1)-(A2.5) describe the production side of the model. To close the model we need to specify the demand for output. We assume that consumers have Cobb-Douglas utility functions and spend a constant share of their income on Q (and the rest on M). The implication of this assumption is that

$$(A2.6) \quad (\dot{Q} - \dot{X}) = -(\dot{p}_q - \dot{p}_x).$$

Equations (3)-(6) determine relative wages and prices and, hence, output and factor allocations. In particular, shifts in the relative employment of college and high school educated labor are related to shifts in relative prices, output and wages, in a linear fashion. Here we concentrate on labor demand:

$$(A2.7) \quad \dot{d} \equiv \dot{L}_c - \dot{L}_h = -\sigma(\dot{W}_c - \dot{W}_h) + \dot{\zeta}.$$

(A2.7) represents the relative demand curve for labor in the state with $\dot{\zeta}$ representing exogenous shocks. The parameter σ represents the elasticity of relative demand between high school and college educated labor. As a result of the fact that we expect $\sigma_x \gg \sigma_q > 1$, we also expect σ to be substantially larger than what we would expect in a closed or large open economy (e.g. the U.S.). The larger are σ_x and σ_q , the larger the share of the workforce in the export sector, and the smaller the discrepancy in factor intensities between sectors, the larger will σ be.

Labor Supply

Supply of college and high-school educated labor is assumed to be the result of two decisions: college education and migration. Education choice is assumed to be a

function of relative wages. The f -curve represents the college flows (relative to high-school educated labor):

$$(A2.8) \quad \dot{f} \equiv \dot{F}_c - \dot{F}_h = \gamma \dot{w} + \dot{\xi} = \gamma(\dot{W}_c - \dot{W}_h) + \dot{\xi} ,$$

where ($\gamma > 0$) is a supply elasticity and $\dot{\xi}$ represent exogenous changes in college flows.

Without migration, (A2.8) would determine the relative supply of college-educated labor. Final labor supply (stock of college educated labor) is determined by the s -curve, which is a flows modified by migration:

$$\begin{aligned} \dot{S}_c &= \theta_c \dot{W}_c + \lambda_c \dot{F}_c \\ \dot{S}_h &= \theta_h \dot{W}_h + \lambda_h \dot{F}_h , \end{aligned}$$

where the $\theta_j > 0$ parameters represent migration elasticities, and the λ_j reflect the fact that net migration into or out of the state may occur at the going wage rate.

Simple case

If we assume that $\theta_c = \theta_h = \theta$ and $\lambda_c = \lambda_h = 1$ (the general case will be addressed later), then relative stocks depend only on relative wages:

$$(A2.9) \quad \dot{s} \equiv \dot{S}_c - \dot{S}_h = \theta(\dot{W}_c - \dot{W}_h) + \dot{f} = (\theta + \gamma)(\dot{W}_c - \dot{W}_h) + \dot{\xi} .$$

Equilibrium

The functions for labor demand, flows, and labor supply for college-educated labor relative to high school labor are described by equations (A2.7)-(A2.9) and generate an equilibrium in relative wages and quantities. Equilibrium changes in wages, flows, and stocks are related to the exogenous shock in the following way:

$$\begin{aligned} \dot{w} &= \frac{\dot{\xi} - \dot{\xi}}{\sigma + \theta + \gamma} , \\ \dot{f} &= \frac{\gamma}{\sigma + \theta + \gamma} \dot{\xi} + \frac{\sigma + \theta}{\sigma + \theta + \gamma} \dot{\xi} , \\ \dot{s} &= \frac{\theta + \gamma}{\sigma + \theta + \gamma} \dot{\xi} + \frac{\sigma}{\sigma + \theta + \gamma} \dot{\xi} . \end{aligned}$$

Positive shifts to flows decrease wages and increase the flow and stock of college-educated labor. The size of the changes in flows and stocks is a positive function of the flow elasticity (γ) and a negative function of the elasticity of relative demand (σ).

Implications for estimation

We are interested in the effect of exogenous supply shifts ($\dot{\xi}$) on stocks, but we do not observe $\dot{\xi}$ directly. Rather, we observe \dot{s} and \dot{f} . Our goal is to estimate the effect of an exogenous shift in flows on stocks, under the assumption of no direct feedback to demand ($\dot{\xi} = 0$). The implied parameter of interest is thus defined as:

$$(A2.10) \quad \beta = \frac{\dot{s} / \dot{\xi}}{\dot{f} / \dot{\xi}} = \frac{\sigma}{\sigma + \theta} \leq 1 ,$$

where β represents the share of the increased flow of college graduates that remain in state in the long run. The effect of an exogenous change in flows on stocks is a function of the demand and supply elasticities. More mobility dampens the effect of flows on stocks through a larger θ . At one extreme, no mobility leads to a one-to-one mapping between changes in flows and changes in stocks ($\beta = 1$ if $\theta = 0$). At the other extreme, frictionless mobility leads to a zero effect of flows on stocks ($\theta \rightarrow \infty$). In contrast, the larger the within state elasticity of demand, the larger will be the effect of changes in flows on stocks. Indeed, as $\sigma \rightarrow \infty$, as is the case in the textbook Heckscher-Ohlin model, $\beta \rightarrow 1$.

The cross-sectional OLS estimate of β is:

$$(A2.11) \quad \hat{\beta} = \frac{Cov(s_j, f_j)}{Var(f_j)} = \frac{\gamma(\theta + \gamma)\sigma_{\xi}^2 + \sigma(\sigma + \theta)\sigma_{\xi}^2 + (2\gamma\sigma + (\gamma + \sigma + \theta)\theta)\sigma_{\xi\xi}}{\gamma^2\sigma_{\xi}^2 + (\sigma + \theta)^2\sigma_{\xi}^2 + (2\gamma\sigma + 2\gamma\theta)\sigma_{\xi\xi}}$$

where there is cross state variation in both ξ and ζ . If $\sigma_{\xi}^2 = 0$, $\hat{\beta} = \frac{\sigma}{\sigma + \theta} = \beta$. If

$\sigma_{\xi}^2 = 0$, $\hat{\beta} = \frac{\theta + \gamma}{\gamma} > 1 \geq \beta$. In general, $\hat{\beta} > \beta$ as long as $\sigma_{\xi\xi} > 0$. This is a sufficient but

not necessary condition for a positive bias, as a zero or small negative correlation still leads to a positive bias.) Since, if anything, we expect some complementarity between the production and use of college-educated labor, we expect $\sigma_{\xi\xi} > 0$. As can be seen, the magnitude of the upward bias in the OLS estimate of β depends positively on σ_{ξ}^2 and negatively on $\sigma_{\xi\xi}^2$, echoing the classic analysis on biases in the estimation of supply and demand equations (Working, 1927).

General case

In the general case where $\theta_c \neq \theta_h$ and $\lambda_c \neq \lambda_h$, we need to work with demand, flow and stock equations for the college and high school educated labor separately. For the more general case, it will be helpful to introduce some new notation: Let E_c and E_h denote earnings share of college and non college graduates and let L_c and L_h denote their population shares.

Demand

Let α represent the export sectors share in total output ($\alpha \equiv p_x X / [p_q Q + p_x X]$), and η the elasticity of demand for exports. Then, the two-sector model developed earlier implies demand functions of the form:

$$(A2.12) \quad \dot{D}_c = -\eta\alpha(E_c \dot{W}_c + E_h \dot{W}_h) - E_h \sigma (\dot{W}_c - \dot{W}_h) + \dot{\zeta}_c$$

$$(A2.13) \quad \dot{D}_h = -\eta\alpha(E_c \dot{W}_c + E_h \dot{W}_h) - E_c \sigma (\dot{W}_c - \dot{W}_h) + \dot{\zeta}_h,$$

where $\eta \geq 0$ is product demand elasticity, σ is elasticity of substitution.

Supply

Flows (supply without migration) are assumed do depend on relative wages. From before, this relationship is: $\dot{f} \equiv \dot{F}_c - \dot{F}_h = \gamma(\dot{W}_c - \dot{W}_h) + \dot{\xi}$ ($\gamma > 0$)

The separate flows can be expressed the following way:

$$(A2.14) \quad \dot{F}_c = \gamma L_h (\dot{W}_c - \dot{W}_h) + L_h \dot{\xi}$$

$$(A2.15) \quad \dot{F}_h = -\gamma L_c (\dot{W}_c - \dot{W}_h) - L_c \dot{\xi}$$

In the stock equations, we relax the assumption of $\theta_c = \theta_h = \theta$ and $\lambda_c = \lambda_h$. In general, we have that:

$$(A2.16) \quad \dot{S}_c = \theta_c \dot{W}_c + \lambda_c \dot{F}_c = (\theta_c + \gamma \lambda_c L_h) \dot{W}_c - \gamma \lambda_c L_h \dot{W}_h + \lambda_c L_h \dot{\xi}$$

$$(A2.17) \quad \dot{S}_h = \theta_h \dot{W}_h + \lambda_h \dot{F}_h = -\gamma \lambda_h L_c \dot{W}_c + (\theta_h + \gamma \lambda_h L_c) \dot{W}_h - \lambda_h L_c \dot{\xi}$$

Equilibrium wages and quantities are the following:

$$\dot{W}_h = c_{hc} \dot{\zeta}_c + c_{hh} \dot{\zeta}_h + c_{hx} \dot{\xi},$$

$$\dot{W}_c = c_{cc} \dot{\zeta}_c + c_{ch} \dot{\zeta}_h + c_{cx} \dot{\xi},$$

$$\dot{f} = \dot{F}_c - \dot{F}_h = f_c \dot{\zeta}_c + f_h \dot{\zeta}_h + f_x \dot{\xi},$$

$$\dot{s} = \dot{S}_c - \dot{S}_h = s_c \dot{\zeta}_c + s_h \dot{\zeta}_h + s_x \dot{\xi},$$

where the c_j, f_j , and s_j parameters are complicated functions of the structural parameters.

Implications for estimation

Using the assumption that demand is not affected directly

$$\beta = \frac{\dot{s} / \dot{\xi}}{\dot{f} / \dot{\xi}} = \frac{s_x}{f_x}$$

For all parameter values we have examined, ceteris paribus, β is decreasing in the migration elasticities (θ_c and θ_h) and the non-economic migration parameters (λ_c and λ_h), and increasing in the demand elasticities (σ and η). The results are available upon request.

The cross-sectional OLS estimate of β will yield:

$$\hat{\beta} = \frac{Cov(s_j, f_j)}{Var(f_j)} = \frac{s_c f_c \sigma_{\zeta_c}^2 + s_h f_h \sigma_{\zeta_h}^2 + s_x f_x \sigma_{\xi}^2 + (s_c f_h + s_h f_c) \sigma_{\zeta_c \zeta_h} + (s_c f_x + s_x f_c) \sigma_{\zeta_c \xi} + (s_h f_x + s_x f_h) \sigma_{\zeta_h \xi}}{f_c^2 \sigma_{\zeta_c}^2 + f_h^2 \sigma_{\zeta_h}^2 + f_x^2 \sigma_{\xi}^2 + 2f_c f_h \sigma_{\zeta_c \zeta_h} + 2f_c f_x \sigma_{\zeta_c \xi} + 2f_h f_x \sigma_{\zeta_h \xi}}$$

We have examined the bias of $\hat{\beta}$ under the assumption of $\zeta_{cj} = -\zeta_{hj}$ (which implies $\sigma_{\zeta_c}^2 = \sigma_{\zeta_h}^2$, $\sigma_{\zeta_c \zeta_h} = -1$, and $\sigma_{\zeta_c \xi} = -\sigma_{\zeta_h \xi}$). The bias is large and positive for all parameter values we examined if the correlation between supply and demand shocks ($\dot{\xi}$ and $\dot{\zeta}$) is nonnegative or small negative. Ceteris paribus, the bias is smaller the larger the variance of supply shocks is relative to the variance of the demand shocks. It is also increasing in

the migration elasticities (θ_c and θ_h), their difference ($\theta_c - \theta_h$), and the non-economic migration parameters (λ_c and λ_h). Demand elasticities have a very small but generally also positive effect on the bias. The results are available upon request.

Appendix 3. Assessing the effects of measurement error in the BA Flow variable

Measurement error may arise because institutional data recorded by year of college graduation rather than year of birth are used in the calculation of the in the BA flow variable. At issue is the magnitude of the bias due to measurement error in BA Flows (number of degrees conferred by state and birth-cohort).

If the age distribution of degrees conferred by year of graduation and state were known, transforming institutional data by year of degree to measures based on birth cohorts would be straightforward. Two problems hinder this strategy. First, state-level measures of the age distribution are not available over a long-time horizon, though national estimates of the age distribution at degree receipt based on calculations from the CPS are available. What is more, since the source of these measures is a sample, these estimated averages are also measured with error. This second problem might be thought of as classical measurement error and, on its own, would lead to attenuation in parameter estimates. However, the overall effect of using national averages on the magnitude and direction of the bias in the parameter estimates is more complicated, depending on both the measurement error caused by sampling variation and the error attributable to the substitution of national measures for state-level variation in the age-distribution at degree receipt. Simulation evidence serves to frame the magnitude of the problem.

Throughout this analysis, we consider the over-time equations of log stocks on log flows; see equation (2) in the text.²³ The introduction of measurement error has a smaller effect on the cross-sectional estimates, because there is substantial stability in state-specific flows across time. As a result the miss-assignment of flows to cohorts is likely to have virtually no effect on our cross-section estimates. On the other hand, such miss-assignment could, potentially, have a large impact on over time estimates.

The Problem

Let F indicate the flow variable measured by degrees conferred and let S indicate the stock variable, the number of people with at least a BA. The level of observation is state, year of birth, and year of measurement for the degrees conferred variable with these levels indicated by subscripts j , g , and t , respectively. The three subscripts are always presented in the order jgt . A dot instead of a particular subscript means sum over all possible values of it. For example, $F_{jg.}$ means total number of degrees conferred in state j among those who were born in year g .

Consider the estimation of effect of flows on stocks, which parallels the dynamic estimates in the text, in the following form:

²³ For notational simplicity, we consider this problem abstracting from the population variables (the denominators of the dependent and independent variables). There is little reason to suspect mismeasurement in the population variables and, in practice, the denominators have very little effect on the estimates.

$$\Delta \ln S_{jgt} = \alpha_{gt} + \beta_t \Delta \ln F_{jgt} + \varepsilon_{jgt}$$

where

$$\Delta \ln S_{jgt} = \ln S_{jgt} - \ln S_{j(g-10)(t-10)},$$

$$\Delta \ln F_{jgt} = \ln F_{jgt} - \ln F_{j(g-10)(t-10)},$$

We want to estimate

$$\beta = \frac{\sum_g Cov_j(\Delta \ln S_{jgt}, \Delta \ln F_{jgt})}{\sum_g Var_j(\Delta \ln F_{jgt})}$$

where Cov_j and Var_j mean covariance and variance across states.

F_{jgt} can be written as the sum of degrees conferred by graduation year (from some minimum age to year t), multiplied by the fraction of those who were born at g :

$$F_{jgt} = \sum_{a=a_0}^{t-g} F_{jg(g+a)} = \sum_{a=a_0}^{t-g} \frac{F_{jg(g+a)}}{F_{j.(g+a)}} F_{j.(g+a)},$$

where a_0 is the lowest possible age of getting BA degree.

Let

$$W_{jgs} \equiv \frac{F_{jgs}}{F_{j.s}}, \quad \bar{W}_{.gs} \equiv \frac{F_{.gs}}{F_{.s}} = \frac{\sum_j F_{jgs}}{\sum_j F_{j.s}}, \quad \delta_{jgs} \equiv W_{jgs} - \bar{W}_{.gs}$$

Then,

$$F_{jgt} = \sum_{a=a_0}^{t-g} F_{jg(g+a)} = \sum_{a=a_0}^{t-g} W_{jg(g+a)} F_{j.(g+a)} = \sum_{a=a_0}^{t-g} (\bar{W}_{.g(g+a)} + \delta_{jg(g+a)}) F_{j.(g+a)},$$

where δ_{jgs} is deviation in state j from the national average $\bar{W}_{.gs}$. Our problem is we observe $F_{j.s}$ and $\bar{W}_{.gs}$ but not δ_{jgs} . Moreover, we measure $\bar{W}_{.gs}$ with a classical error $\omega_{gs} \sim iidN(0, \sigma^2_{\omega})$ because it comes from survey samples (CPS). The variable we use for estimating our model is

$$\tilde{F}_{jgt} \equiv \sum_{a=a_0}^{t-g} (\bar{W}_{.g(g+a)} + \omega_{.g(g+a)}) F_{j.(g+a)}$$

So that instead of β we end up estimating

$$\tilde{\beta} = \frac{\sum_g Cov_j(\Delta \ln S_{jgt}, \Delta \ln \tilde{F}_{jgt})}{\sum_g Var_j(\Delta \ln \tilde{F}_{jgt})}.$$

Simulation strategy

The results of simulations indicate that while ω increases the variance of the right-

hand side variable in the classical way, using $\bar{W}_{.gs}$ instead of $\bar{W}_{.gs} + \delta_{.gs}$ actually decreases it. The latter effect, however, depends largely on the time-series properties of δ : the more serially correlated it is the smaller the artificial decrease in the variance. Moreover, δ may well be correlated with either F or S (across states), which induces further complications.

The sizes of the CPS samples provide good estimates for σ^2_{ω} . Therefore, we can carry out simulations once we specify a data generating process for δ and calibrate its parameters. Two aspects are important: serial correlation and cross-state correlation with the F and S series. We have census-based estimates for W_{jg1970} and W_{jg1980} . Assuming that those are true values without measurement error (the Census samples are very large), we can calculate δ_{jg1970} and δ_{jg1980} in such a way that preserves their actual cross-state correlation with F and S for those years. Unfortunately, two points in time do not convey much information about the time-series properties of a series. A further complication arises because δ has to satisfy some additivity constraints. Moreover, it turns out that neither fitting a linear trend on the two points nor fitting an AR(1) process are good specifications because they yield nonsensical values for the 1950s. Instead, we have specified several data generating processes through the available points of δ_{jg1970} and δ_{jg1980} .

Results from two main specifications are presented below. In specification 1, δ is fixed to its 1980 value. In specification 2, δ follows a piecewise linear trend: it is fixed to its 1980 value in 1980 and after, fixed to its 1970 value in 1970 and before, and changes in a linear fashion between 1970 and 1980. We have experimented with adding *iid* disturbances to the trends (the standard error was calibrated to be half of the absolute deviations between the 1980 and 1970 values). They gave results identical to the deterministic specifications up to the fourth decimal point. Also, we have experimented with fixing δ to its 1970 value instead of 1980. The results from these two were essentially the same. Throughout the exercise, ω was calibrated as an *iid* Normal random variable, with variance $\bar{W}(1-\bar{W})/(sample\ size)$.

In each simulation run, we generated δ and ω series, simulated the “true” W using the available national average and δ and ω , and from W and the graduation year-specific flows, the “true” birth cohort specific flow variables. Then we estimated the model. Then we took the national average of the simulated “true” W variables, generated the flow variable using the \bar{W} , and estimated the model again. These two estimates were then compared to get the bias. Table A3 shows the results.

Our estimate of the bias arising from the measurement error in the national average estimate is always negative, as expected, and is between 1 and 5 percentage points. Our estimate of the bias arising from using the national average rather than state specific age distributions varies by year, age group, and the method we use to impute δ 's for the intercensal years. More often than not, our simulations suggest that the use of the national average data tends to inflate rather than attenuate estimates. The estimated bias

due to the national average varies between -3 and $+13$ percentage points. The result of the two sources is a bias most likely to be between -5 and $+10$ percentage points. More often than not, our simulations suggest that the use of the national average data tends to inflate rather than attenuate estimates. There is no suggestion from this simulation evidence that our OLS estimates are severely attenuated.

Table A3.1. Simulation results: average bias of β in the over time models.

Age	Year	Point Estimates ^a	Average bias (biased minus simulated “true” estimate)					
			Specification 1 ^b			Specification 2 ^c		
			Total	Due to ω^d	Due to \bar{W}^d	Total	Due to ω^d	Due to \bar{W}^d
25-34	1960 - 70	0.21	0.00	-0.04	0.04	-0.04	-0.04	0.00
25-34	1970 - 80	0.10	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
25-34	1980 - 90	0.60	0.04	-0.05	0.09	0.07	-0.05	0.13
35-44	1970 - 80	0.26	-0.01	-0.02	0.01	0.10	-0.02	0.12
35-44	1980 - 90	0.20	-0.05	-0.02	-0.03	0.09	-0.02	0.11
45-54	1980 - 90	0.06	0.00	-0.01	0.01	0.02	-0.01	0.03

The figures are average values from 100 Monte Carlo replications.

^a The point estimates are the biased estimates from the specification used in the simulations. They are slightly different from the original estimates in Table 5 because the latter are based on variables divided by population.

^b Specification 1: δ is fixed to its 1980 value; $\omega \sim iidN(0, \sigma_\omega^2)$, $\sigma_\omega^2 = \bar{W} (1 - \bar{W}) / (\text{sample size})$. For the definition of δ and ω see the text above.

^c Specification 2: δ is fixed to its 1980 value in 1980 and after, fixed to its 1970 value in 1970 and before, and changes in a linear fashion between 1970 and 1980.; $\omega \sim iidN(0, \sigma_\omega^2)$, $\sigma_\omega^2 = \bar{W} (1 - \bar{W}) / (\text{sample size})$.

^d “Due to \bar{W} ” means bias due to using national averages instead of the state-specific age distributions and is based on the assumption that there is no measurement error in \bar{W} (results from runs with $\omega = 0$ everywhere). “Due to ω ” means bias due to sampling error in the measures national averages and is measured as “Total” minus “Due to \bar{W} ”.