Attrition from Administrative Data: Problems and Solutions with an Application to Higher Education[*]

Andrew Foote
U.S. Census Bureau
Andrew.Foote@census.gov

Kevin Stange
University of Michigan and NBER
kstange@umich.edu

November 2019

ABSTRACT

Recent research in many fields of social science makes extensive use of administrative data, such as from US states, counties, or school districts. Recent work on the labor market consequences of post-secondary education, in particular, have used administrative data from institutions matched to in-state earnings data. However, few of these papers have the ability to follow workers outside of the state, which could bias measured effects on earnings. Similar problems arise outside the US, when workers migrate across countries. While most researchers acknowledge the issue, they are unable to quantify the effect that non-random attrition has on their results. In addition to these academic papers, a number of states and countries have produced and publicized average earnings of graduates to inform students. Using new data merging college records with both in-state and national earnings from the LEHD, this paper documents how earnings estimates are biased in practice. We also document how this differs by field of study and college selectivity, as well as the extent to which attrition is differential across the earnings distribution. We find that out-of-state migration is particularly problematic for high-earners, flagship graduates, and business majors and grows with time since graduation. In our empirical example, we find that the effect of graduating from a flagship university (relative to less selective public 4-year) is 26% higher than one would estimate using in-state earnings exclusively. Earnings differences across majors are also understated. Various approaches to testing for and bounding this bias are considered.

[*] Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been approved for disclosure with DRB Requests #DRB-B0007-CED-20181029, #DRB-B0033-CED-20190318, #DRB-B0064-CED-20190703

I. Introduction

There has been a big shift from the use of survey data to administrative data in social science generally (Penner & Dodge, 2019) and economics and education in particular (Chetty 2012; Figlio et al 2017). Relative to survey data, administrative data has several benefits including larger samples, lower cost, less measurement error and more extensive longitudinal follow-up (Card, Chetty, Feldstein, Saez, 2010; Figlio et al 2017). An important aspect of this longitudinal follow-up is that administrative data generally have much higher response rates and lower attrition than survey data since inclusion is generally not an active decision that sample participants must make. Prior work has found that survey non-response and attrition can generate considerable bias (Lillard, Smith and Welch, 1986; Bollinger, Hirsch, Hokayem, Ziliak, 2018), as can sample selection more generally (Lee, 2009). Many view the administrative data revolution as eliminating, or at least mitigating, much of this non-response and attrition bias.

However, much of the administrative data used comes from administrative units at a subnational level, such as cities, school districts, counties, and states. Researchers have used such sources to study educational outcomes, criminal justice outcomes, health care utilization, earnings, and participation in social insurance programs such as SNAP or unemployment insurance. The use of such subnational administrative data can introduce bias if study participants are mobile across jurisdictions.[1] For instance, researchers can usually not distinguish whether a study participant absent from state administrative earnings data is truly not working or is working in another state. The same problem potentially arises even when using national administrative data (e.g. tax records); many OECD countries have double-digit emigration rates for high-skilled workers (OECD, 2015)

In this paper we examine the bias arising when using administrative data to measure program outcomes in the presence of attrition. We illustrate the problem with the case of estimating the labor market effects of college quality and college major, a growing literature that has made extensive use of state-level administrative earnings data collected to administer unemployment insurance. The key challenge is that earnings is not observed if participants move out of state, so

[1] Merging administrative data across domains can also potentially introduce bias from cases that cannot be uniquely merged for various reasons, such as name changes, misspellings, or lack of unique identifiers. This source of bias is also potentially present when using national administrative data sources. We ignore this issue in this paper.

researchers are unable to distinguish non-employment from interstate mobility. This is problematic if migration is affected by the treatment under study, which is likely given that migration differs with college attainment (Malamud and Wozniak, 2012), with financial aid receipt (Fitzpatrick and Jones, 2016), and across majors (Ransom, 2016). Our analysis is made possible by a special link between education records from twenty-three universities in Texas and Colorado and the U.S. Census Longitudinal Employer-Household Dynamics (LEHD) data, which combines UI earnings records from all states and the District of Columbia. Thus we can validate analyses that use in-state (subnational) earnings records to those using national records. This novel data allows us to answer four questions: (1) How significant is out-of-state migration for recent college graduates? (2) Does migration differ across the earnings distribution? (3) Does it impact estimates of earnings effects of college selectivity and major? (4) What should researchers do about it?

Prior work has either focused on workers with in-state earnings (Hoekstra, 2010; Andrew, Li, Lovenheim, 2016; Andrews and Stange, 2019; Altonji and Zimmerman 2018) or set non-matched workers as having zero earnings, often in conjunction with a bounding exercise (Denning, Marx, Turner, 2018). These researchers have not been able to directly test the validity of these approaches. Our work is most directly related to two studies that assess migration-related bias from using UI administrative data. Scott-Clayton and Wen (2017) use the NLSY97 to demonstrate how estimates of the earnings effect of college attainment are affected when using only earnings records for students that remain in state. They also find that out-of-state migration tends to attenuate the earnings premium of a college degree. When constructing institution-specific earnings outcomes contained in the College Scorecard, Council of Economic Advisors (2017) compare estimates derived from the full universe of IRS tax records with those using in-state employee records only, to approximate the restrictions of state UI earnings records. They find that migration bias overstates the average earnings of graduates from low-earning colleges, understates that from high-earning colleges, and is larger when out-migration rates are higher. Our paper complements these two by examining a much larger administrative data sample than the NLSY and by using the state administrative data that most researchers and states have access to (in contrast to IRS tax data). We also begin to assess several tests and corrections that researchers have proposed to address the problem.

Our approach is also in the spirit of work that validates survey with administrative data. Bollinger Hirsch Hokayem and Ziliak (2018) validate CPS earnings variables using Social Security administrative records, given the large non-response in the former. Barnow and Greenberg (2015) compare various social experiments using both survey and administrative earnings data. Britton, Shephard, and Vignoles (2018) compare labor market outcomes in the UK Labour Force Survey to administrative records, finding substantial differences between the two. The differences result in different conclusions about important labor market phenomenon, including the gender wage gap, the returns to education, and the extent of earnings inequality. We also contribute to the broader literature proposing solutions to various forms of selection bias (Lee, 2008), attrition bias (Grogger, 2013), and non-response bias (Behaghel, Crepon, Gurgand, LeBarbanchon (2015; 2009).

We find that migration out of state is considerable, approaching 30% even among graduates attached to the labor force. Furthermore, it is not ignorable, as mobility is higher for students at the higher end of the earnings distribution, for certain majors, and for certain institutions. Monte Carlo analysis suggests that a key factor is the relative treatment effect on out-of-state vs. in-state earnings. Migration-related bias will be negative if attending a selective institution or graduating with a STEM major increases out-of-state earnings more than in-state earnings. Surprisingly, this is true even if migration is exogenous. Bias is zero when treatment has a similar effect on in-state and out-of-state earnings, even if migration is endogenous.

Migration also increases with time since labor force entry, so long-term follow up will be more subject to migration bias than short-term follow-up. This is problematic, as longer-term earnings outcomes are more pertinent to welfare and are pushed in use of performance measures by states because they are more stable (Miniya and Scott Clayton, 2018).

We illustrate the practical consequences of this migration-induced bias for estimates of the earnings differences between flagship and non-flagship graduates. Flagship graduates who are working earn $3465 (11 log points) more per quarter than non-flagship graduates, though in-state earnings records would understate this premium by 26%. The inclusion of modest controls for selection into a flagship does not mitigate this problem. The magnitude of the bias is greater at higher points on the earnings distribution and with greater time since graduation, reflecting the greater rates of differential migration among flagship graduates along those two dimensions.

Interestingly, our bias results appear to be driven by bias in flagship earnings for Colorado but not Texas, although the differences in mobility between flagship and non-flagship graduates are similar in the two states. We discuss the implications of this finding in light of empirical tests that researchers perform to test for bias. Earnings differences across fields can also be substantially under-stated by the use of in-state earnings records.

We offer three practical lessons for researchers using such data. First, bias is reduced when the sample is conditioned on having positive observed earnings. Doing so changes the target of estimation to a parameter that does not fully capture the consequences of the treatment under study (and could be subject to standard Heckman-like selection into working), but this drawback may be the lesser of two evils compared to erroneously assuming movers are not working. Second, the Lee (2007) bounding approach is likely inappropriate in this setting given the failure of the monotonicity assumption and bounding approaches that relax that assumption are wide and uninformative. Third, a common test of the potential of bias is whether the rate of in-state earnings being observed differs between treatment and control groups. While the potential for bias is certainly larger when attrition differs between treatment and control, this test is not definitive of the presence of bias. We also show in our setting that similar levels of missing-ness can produce quite different levels of earnings bias. Our Monte Carlo analysis shows that even having migration unrelated to treatment is not sufficient to rule out bias. Nor is a relationship between moving and actual earnings necessarily evidence of bias. They key factor is whether migration is correlated with earnings differentially with the treatment understudy. This is inherently not testable, though supplemental data might be suggestive if available. Ongoing work is assessing other approaches for testing and correcting for the bias.

Our analysis also has two implications for states. First, states do not retain many of their highest-paid workers, which is a goal of many state merit-aid programs. Second, earnings estimates published by state higher education boards and made available to students will understate earnings differences between programs (institutions and fields) due to systematic differences in rates of out-migration. Published earnings records for smaller states with high rates of out-migration will be particularly misleading.[2]

This paper proceeds as follows. In the next section, we provide a selective review of recent work that uses administrative records to estimate treatment effects. We focus on earnings outcomes used to measure the effect of postsecondary choices and treatments. Section III describes our data and samples. Section IV presents our empirical results, including descriptive evidence on cross-state migration by selectivity, major, and time since degree and regression analysis of the effect of flagship graduation. In Section V we discuss the tests and bounding approaches used in the literature and evaluate the performance of common bounding techniques. Evaluation of other tests and corrections mentioned in this section is ongoing. In Section VI we describe Monte Carlo simulations of a simple model that illuminates the conditions that give rise to biased estimates of treatment effects, permitting us to speak to settings more general than our specific empirical example. Finally, section VII concludes.

II. The Use of Sub-National Administrative Earnings Data

Administrative earnings data has been used extensively by researchers to study the effects of various choices and treatments in higher education.[3] Table 1 lists numerous recent examples from both the US and international contexts. Such data has permitted researchers to estimate the labor market effects of college quality (Hoekstra, 2010; Andrews, Li, Lovenheim; 2016; Minaya and Scott-Clayton, 2018; Cunha and Miller, 2014), college attendance (Zimmerman, 2014; Turner, 2014; Ost, Pan, Webber, 2018); degrees (Jepsen, Troske, Coomes, 2014; Engborn and Moser, 2017), and major or program of study (Bakkes, Holzer, Valez, 2015; Stevens, Kurlaender, Grosz, 2018; Altonji and Zimmerman, 2018; Andrews and Stange, 2019). Outside the US, researchers have been able to exploit institutional features that let them credibly estimate the earnings effects of field and program (Hastings, Neilson, Zimmerman, 2013; Kirkeboen, Leuven, Mogstad, 2016; Belfield et al, 2018).

[2] One caveat is that states may be particularly interested in the earnings of graduates that remain in state, since this has important implications for tax revenue. The earnings of graduates who leave the state may be less relevant in this setting.

[3] We focus here on work related to higher education, but examples outside of higher education are numerous. For instance, recent studies of displaced workers (Lachowska, Mas, & Woodbury, 2018), housing demolitions (Chyn, 2018), incarceration, (Mueller-Smith, 2018), and foster care (Doyle, 2013) all use administrative data from one state to measure outcomes.

Furthermore, many U.S. states and postsecondary systems have begun publishing interactive tools that allow students to see the consequences of college major choices. For example, these tools are available for at least a subset of institutions in Florida, Texas, Virginia, Colorado, Iowa and Tennessee. Some of these tools release earnings by major, while others focus only on outcomes by institution. All of these tools use matched post-secondary and workforce records of program graduates who stay in state.[4]

One thing these studies have in common is that estimates could be subject to bias due to migration out of the jurisdiction for which outcomes are captured, be that a state or country. Table 1 also reports out-migration rates by state and emigration rates by country. The overall five-year cross-state migration rate in the U.S. is approximately 9%, though this is likely higher for young college graduates. Furthermore, there is quite a bit of variation across states, suggesting that the potential for bias likely differs across states. Though not really comparable to the US cross-state migration figure, rates of emigration from OECD countries are also high, particularly for high-skilled workers.

Authors in these papers have taken several approaches to address the potential sample selection problem. Most studies focus on workers with in-state earnings (Andrew, Li, Lovenheim, 2016; Hoekstra, 2010; Altonji and Zimmerman 2018; Andrews and Stange, 2019), dropping workers with no in-state earnings over some time frame. This approach assumes that dropped workers are similar to non-dropped workers. Other papers retain non-matched workers, setting their earnings to zero, often in conjunction with a bounding exercise (Denning, Marx, Turner, 2018). Many studies test whether treatment is correlated with having matched outcome data, interpreting no effect as evidence of minimal bias. While appropriate if selection is one-directional, we will see that this test does not rule out substantial bias if treatment induces selection on multiple margins. Few researchers have explicitly examined whether treatment is related to the probability of being in-state or directly looked at inter-state migration using other sources (e.g. ACS). One exception is Andrew, Li, and Lovenheim (2016), who compare the earnings distribution of recent college graduates that are living in Texas vs. out of Texas among those who lived in the same college

[4] An effort underway at the U.S. Census Bureau provides estimates of earnings by institution, degree level, and degree field. To see more, see: https://lehd.ces.census.gov/data/pseo_beta.html

town five years earlier from the 2000 Census. While suggestive of minimal bias, this test is not conclusive and not possible for many treatments under study.[5]

III. Data Sources for Empirical Application

We examine this issue using new data linkages at the U.S. Census Bureau between university transcript records and a national database of employment and earnings.[6] Our analysis includes enrollment and graduation data for students from the University of Texas System and all public universities in the State of Colorado.[7] These data includes degree field, graduation date, degree level, and data on subsequent enrollments. A current limitation of the data is that it contains very few baseline demographic variables. To complement the administrative data from the Texas and Colorado, we also use administrative data from Census to maintain consistent demographic data for our sample. We restrict our analysis to baccalaureate graduates, since much of the research on labor market outcomes of graduates has focused on this population. These two systems include nine campuses in Texas and fourteen campuses in Colorado, which collectively span a wide range of institutional size, selectivity, and resources, as reported in Appendix Table A1. We primarily focus on comparisons between students graduating from the two flagship institutions (University of Texas at Austin and University of Colorado at Boulder) vs. other public four-year institutions in the two states.

Student records are matched to the Longitudinal Employer-Household Dynamics (LEHD) data. LEHD data reports quarterly earnings by job (employer-employee match) for all employment covered by unemployment insurance, including those on paid leave. These data do not include the self-employed (independent contractors and unincorporated self-employed), railroad workers covered by the railroad unemployment insurance system, and some smaller categories of workers (some family employees, certain farm workers, etc). Most state and local government

[5] In the presence of swapping, this test may be biased toward not finding any differential effect, and is therefore underpowered. Furthermore, the geographic location of many flagship campuses are not separately identified in the public use versions of the Census or ACS.

[6] These data linkages are part of a larger project, which has included the creation of the experimental data product Post-Secondary Employment Outcomes. See https://lehd.ces.census.gov/data/pseo_beta.html for a description of the project as well as the tabulations that have already been released. Technical documentation is available at https://lehd.ces.census.gov/doc/PSEOTechnicalDocumentation.pdf

[7] These data include in-state resident and out-of-state students; we do some analysis disaggregating them, and plan on doing more in that respect.

employment is included. These data span 2000-2016 for 50 states and the District of Columbia. The LEHD data cover approximately 96% of all private sector employment, though the overall coverage of all employment (including self-employment, all public sector, etc) is lower (Abowd et al., 2009). We ignore this incomplete coverage and assume that any individuals not matched to the LEHD have zero earnings nationally (and in-state). Importantly, this data allow us to measure earnings for graduates that leave the state, which is the main contribution of our paper. For each graduate, we calculate national and in-state annual earnings separately for each quarter since graduation, in order to measure the bias of only measuring in-state earnings. Importantly, our national and in-state earnings measures come from the same source (with in-state a subset of the national), so any difference can be attributed to differences in coverage, not variable definition. All earnings amounts are converted to real 2018 dollars using the CPI-U.

Our full analysis sample comprises a 10 percent sample of students who graduated from one of these twenty-three campuses from 2001 to 2013, though in some analyses we restrict to graduates from 2006 and earlier so that we can have a balanced panel of individuals when looking at earnings outcomes over different time horizons. Each observation is a person-quarter, beginning with the first quarter of the first calendar year after graduation and going up to 15 years post-graduation for our earliest cohort. Our full analysis sample includes more than half a million earnings observations for flagship graduates and 1.2 million observations for non-flagship graduates.

Table 2 presents summary statistics, separately for students graduating from one of the two flagships and all others. Looking at total national (true) earnings and pooling all quarters since graduation, the earnings advantage of flagship graduates is apparent: flagship graduates earn $2,500 (22%) more in quarterly earnings than non-flagship graduates. However, erroneously treating migrants as having zero earnings by only looking at in-state earnings, flagship graduates appear to earn $150 less than non-flagship graduates. These differences arise because rates of non-employment and out-of-state migration differ between flagship and non-flagship graduates, as shown in the final rows. Flagship graduates are about 10 percentage points more likely to have moved and worked out-of-state than non-flagship graduates, with the gap increasing with time since degree. Interestingly, flagship graduates are actually slightly *less* likely to have any positive earnings nationally.

IV. Results

A. Graphical Evidence: Is Migration Ignorable?

To set up our regression analysis, we first establish several facts about the migration of college graduates using graphical evidence. Collectively, these patterns suggest that missing earnings data could affect empirical treatment effect estimates. We use institutions from Colorado and at the University of Texas System.[8] We should note that Texas has a relatively low out-migration rate of young workers (6.7% vs. 8.7% for the U.S. overall, as reported in Table 1), suggesting the problem we illustrate may be even more pronounced in other states, while Colorado has relatively high rates of out-migration.

First, graduates (from all institutions) leave the state at appreciable rates. Figure 1 shows the share of workers attached to the labor force that stay in Texas by year after graduation. In this graphical analysis we restrict the sample to graduates who have at least three quarters of non-zero earnings and with at least $10,000 of earnings in the calendar year nationally (in-state or out-of-state). These restrictions are intended to capture, in an imperfect way, people that have reasonable attachment to the labor market earning at least the minimum wage. In the first year after graduation, almost 80 percent of graduates are employed in Texas. However, that number falls steadily and appreciably through ten years after graduation, when fewer than 70 percent of employed graduates are employed in Texas.

Second, those that leave the state have measurably different earnings than those that stay in the state. If the stayers have similar outcomes as the movers, then migration is ignorable. To assess this, in Figure 2 we plot the share in-state by individuals' location on the national earnings distribution among graduates from our institutions, separately for 1, 5 and 10 years after graduation. Note that these graphs impose the same earnings restrictions as above. At all time periods, migration has a "U-shape" relationship with earnings: low and particularly high earners are more likely to move out-of-state than middle-earners. Migration is clearly related to earnings, though in a non-linear way. While Figure 1 shows that the share of graduates that stay in-state is steadily falling, Figure 2 shows that the downward shift in the share in-state is not constant across the earnings distribution. Instead, those that leave the state are much more likely to be in

[8] The caveat to this is that Figure 1 only includes the University of Texas schools.

the higher earnings percentiles, and that is especially true 10 years after graduation. Graduates at the 50th percentile of the national earnings distribution were about 10 percentage points less likely to be in state after ten years; graduates at the 90th percentile were over 15 percentage points less likely. Out-migration rates are approaching 50% after ten years for the highest earners.

Because high earnings may be associated with mobility by construction (those who leave are high earners due to cost of living differences, for example), we also created the same graph, but fixed income for individuals by constructing a "lifetime earnings" measure, which sums up all earnings in the first 10 years, and graph the resulting mobility by percentile earnings in Figure 3. This figure shows that mobility is higher for higher earners, but it is particularly higher at the tail of the distribution.[9]

Figures 1 and 2 taken together illustrate an important issue in studying long-term earnings outcomes for students when restricted to one state. In-state earnings are a better proxy for total national earnings in years immediately following graduation. However, as Minaya and Scott-Clayton (2016) argue, early earnings years are very noisy and are unlikely to accurately measure the true effect of a postsecondary treatment, such as attending a specific college. If instead researchers focus on later earnings years for graduates who stay in-state, they are likely capturing a biased estimate of the treatment effect because they do not measure the effect for mobile workres.

Third (and most relevant to studies of the effect of college quality) is that migration patterns differ by institution. Figure 4 shows the same estimates of share in-state, except this time by institution. Panels (a) and (c) compare the in-state share for three UT schools one and ten years after graduation: two large state schools and the flagship (UT Austin). Similarly, panels (b) and (d) compare in-state share for three Colorado schools one and ten years after graduation: Colorado State University, University of Colorado-Boulder (the flagship) and Colorado School of Mines (an engineering school). The graphs illustrate two important points. First, graduates of the flagship are much more likely to move at all points in the distribution. The exception to this

[9] One explanation for the similarities across time periods for this sample is that it is highly selected, including only people who work for at least three quarters in each year and earn $10,000 per year for each year of the ten years after college.

rule is the comparison with Colorado School of Mines, which is not a flagship, but is a nationally-ranked engineering school. Second, there are differences in attrition by state. While UT Austin has higher attrition rates overall, there are not differences across the earnings distribution, while CU Boulder has much higher rates of attrition at the top end of the earnings distribution. We return to these differences between the mobility of flagship graduates in the next section when discussing the earnings bias of estimating returns to flagship graduation. This result suggests that work done to measure the effect of college quality on earnings, if restricted to in-state earnings, may substantially underestimate the returns to college quality, if the missing data is disproportionately drawn from the upper tail of the earnings distribution.

We also look at other margins of mobility across the earnings distribution. Figure 5 shows in-state share for the top 5 majors at the universities in our sample one year after graduation, and illustrates that there are considerable differences in mobility across the distribution by major field. Health and engineering majors are generally equally likely to move regardless of income, while mobility is much higher at the higher end of the earnings distribution for business and social science majors.

Figure 6 shows in-state share by gender for one and ten years post-graduation. While mobility is similar in year 1 for the bottom half of the earnings distribution, males are much more mobile at the top of the distribution. By 10 years after graduation, there is attrition at the top of the distribution for both genders, but it is more pronounced for males. Many studies focus (implicitly or explicitly) on students who were previously in-state residents. Figure 7 shows that while rates of mobility are lower for in-state resident students than out-of-state students, the difference between mobility at the 50^{th} percentile and 90^{th} percentile is pretty similar – a drop of about 7 percentage points. Additionally, at 10 years after graduation, the mobility is *more* differential at the top of the earnings distribution for in-state students, while out-of-state students move at high rates at every point in the earnings distribution.

To summarize, it is clear that the earnings outcomes of graduates leaving the state are measurably different on a number of dimensions, creating conditions for migration-related bias when estimating postsecondary treatment effects using in-state administrative data.

B. Bias in Effect of Flagship

The graphical analysis above illustrated three ingredients for biased estimates of the effect of flagship graduation on earnings: (1) substantial out-of-state migration; (2) migration patterns that differ by position on the earnings distribution; and (3) a differential earnings-migration link across institutions. Furthermore, these patterns all differed with time-since-degree, suggesting that the extent of bias will differ over different time horizons, with later earnings measures more susceptible to bias. We now evaluate how these factors contribute to bias estimates of the effect of graduating from a flagship university. Our regression analysis uses the complete sample from both Colorado and Texas, summarized in Table 2. We pool across all quarters since college graduation.

We estimate simple OLS regression models of observed earnings on a dummy for having graduated from a flagship university:

$$Y_{it} = \beta_0 + \beta_1 Flagship_i + \beta_2 X_{it} + \varepsilon_{it} \tag{1}$$

Since we are combining many quarters and cohorts, we include a full set of calendar year, quarter, and graduation year dummies (though not the interactions) to control for any lifecycle and cross-cohort earnings trends that may happen to correlate with flagship enrollment. We also include a state fixed effect to account for the fact that earnings and flagship enrollment may differ between states. We have limited demographic controls, but we also include sex and race dummies in some models. Standard errors are clustered at the individual (person) level.

Our data permits us to estimate models using true earnings from all sources nationally and identical models using outcomes constructed from the in-state data typically available to researchers. Non-matched records are set to zero and are included or excluded depending on the specification. Our empirical construct of bias is simply the difference between these two estimates.[10] We should note that due to the limited number of control variables and a lack of plausibly exogenous variation in flagship attendance, we do not interpret our estimates as the causal effect of flagship enrollment. However, we use the terminology of "effect" to be consistent with the treatment effect literature. The migration bias problem we describe is not

[10] This difference also includes estimation error. In the future we will construct standard errors for the bias that accounts for this estimation error.

mitigated by having exogenous variation in treatment.[11] However, since different identification strategies may estimate treatment effects for different local populations (with different rates of differential migration), the extent of bias could differ across methods.

Table 3 presents our results. Panel A presents the most naïve estimates: earnings differences in levels including any non-matched records as zeros. As was apparent from the summary statistics, doing so causes large negative bias. Relying on in-state records only would lead a researcher to conclude that flagship graduates earn \$700 less than non-flagship graduates whereas they actually earn more than \$2000 more per quarter. The resulting bias is substantial: any effect of flagship on migration instead appears as a large reduction in earnings (to zero).

Recognizing the potential for this bias, most scholars have instead focused on individuals with some attachment to the labor market, as indicated by having non-zero earnings.[12] Panel B restricts to quarterly observations with non-zero earnings either in-state (columns 1 and 4) or nationally (columns 2 and 5).[13] Doing so greatly reduces the bias because any effect of treatment on out-migration rates is no longer recorded as a large reduction in earnings. Nonetheless, doing so does not eliminate the bias. In-state earnings records will understate the true flagship effect by \$719 per quarter; the true flagship effect is 26% higher than the in-state earnings records would suggest. This finding is similar if the log of quarterly earnings is used as the outcome, as is commonly done. In-state earnings would suggest a 0.087 log point premium to flagship enrollment, whereas the true effect is 0.024 log points (27%) higher. Andrews, Li, and Lovenheim (2014) find that UT-Austin graduation increases earnings more at the high end of the distribution, suggesting that "this university is particularly lucrative for top earners." The last four rows present quantile regression estimates of the effect of flagship graduation on various moments of the earnings distribution. We too find that the flagship earnings premium increases across the earnings distribution, from 6% at the 90th percentile up to 17% at the 90th, using in-

[11] We show this later in the selection model, and it is also apparent in our Monte Carlo simulations in which treatment is randomly assigned.
[12] Some researchers have restricted it to quarters with non-zero earnings whereas others have restricted it to years with several quarters of non-zero earnings or earnings greater than some minimal threshold.
[13] We should note that imposing this restriction creates the well-known sample selection problem (Heckman, 1974): earnings outcomes for those observed to be working will be different than those choosing not to work. We abstract from this issue, treating the self-selected national earnings outcome as our target for estimation.

state earnings. However, the magnitude of the bias is also growing across the distribution, from 0.017 log points at the 25th percentile up to 0.024 log points at the 90th.[14]

We also ran the results separately for Texas and Colorado with log earnings as the outcome variable, and find that the bias findings are driven entirely by Colorado (bias of 5.6 pp), while Texas has effectively no bias (bias of -0.004pp). We also summarize these results graphically in Figure 8, which plots the estimated bias for Colorado and Texas separately by year post-graduation. It shows that Texas estimates consistently have zero bias, while Colorado has large bias. We have two hypotheses that may explain these results. First, the Texas sample could differ from Colorado in important ways because it includes a more select group of universities (just the University of Texas schools), which may have different levels of differential attrition than less selective institutions.[15] Second, it may be that bias is worse in the presence of higher overall migration, since Colorado has much higher out-migration rates. However, the difference in migration between non-flagship and flagship institutions is very similar in the two states (0.162 for CO and 0.160 for TX), as shown in Panel B.

Importantly, our analysis seems to confirm that the test by Andrews, Li and Lovenheim (2016) was informative, in that there was no differential migration across the earnings distribution for the flagship institution. However, many papers cite their test as confirmation that there is likely no bias due to migration more generally (not just for flagship institutions in Texas), which we have shown is not the case.[16]

We lack rich controls and credible quasi-experimental variation to estimate convincing causal effects of flagship graduation on labor market outcomes. So the question arises of whether such a setting would be less susceptible to the migration bias we uncover. The third set of columns in Table 3 include controls for sex and race (4 categories) as suggestive evidence on this question. While including these controls reduces the flagship premium by several percentage points, there

[14] The proportionate size of the bias is actually decreasing across the earnings distribution, from 30% at the 25th percentile to 13.5% at the 90th because the base premium is lower at lower percentiles.
[15] Census is in activity negotiations with THECB, so we can eventually include all universities in Texas in our analysis.
[16] A number of recent papers directly cite this result from Andrews et al. (2016), including Denning, Marx and Turner (2019) and Andrews and Stange (2019).

is no impact on the migration bias. Magnitudes of the bias are quite similar or even larger in some cases and the bias as a proportion of the true effect is even larger than with no controls.

Table 4 reports results separately for different years since graduation. The general pattern of substantial bias, particularly when including non-matched records as zeros, is true at all time horizons. Furthermore, the magnitude of the bias generally increases with time since graduation overall and at most points on the earnings distribution, consistent with the increased rate of migration rate of flagship graduates over time.[17] The implication is that though longer-term earnings may be a better proxy for welfare consequences of a treatment, these outcomes are more susceptible to migration-related bias when using in-state earnings outcomes.

C. Majors

While our main application is to earnings differences between flagship and non-flagship institutions, the issue may also present itself in estimates of outcome differences across major fields, given the differential migration across fields, which we show above in Figure 5. (see also Ransom, 2016) Cross-major differences in bias are important given that several states are now publishing program-level labor market outcomes using in-state earnings data. We demonstrate the bias in earnings at the major level in two ways. First, we plot the difference in average earnings between in-state and national earnings outcomes. Figure 9 demonstrates the potential for bias in earnings estimates by major. For each major, we calculate the average annual earnings using in-state and national earnings data and then calculate the difference in these averages by year since graduation.[18] Each panel plots these differences for all broad fields of study (2-digit CIP), but highlights one of four specific major fields. Importantly, each major demonstrates a different level and time path of national vs. in-state earnings. In-state earnings substantially understates the average earnings of Computer Science majors, with particularly large differences after six years. On the other hand, national and in-state earnings for psychology majors are pretty close, implying minimal bias. The difference for business majors grows with time since graduation, whereas the difference for engineering majors is fairly stable. More generally, cross-

[17] One slightly puzzling finding is the negative flagship effect in year 1. This appears to be due to the greater rate of inclusion of observations with very low earnings in the first year at the flagship, which could be due to greater rates of graduate school enrollment among flagship graduates.
[18] For these graphs, the sample includes all graduates from our Texas institutions who have non-zero annual national earnings.

major gaps in earnings coverage are relatively small immediately after graduation, but spread out substantially by ten years. This will result in biased estimate of the treatment effect of specific majors, such as CS relative to Psychology, with the extent of bias evolving (likely growing) with time since graduation.

The second way that we illustrate the bias of in-state earnings outcomes by major is by estimating major-specific fixed-effects in two sets of regressions:

$$Y_{it}^{in} = \beta_0 + \gamma_f^{in} + \beta_1 X_{it} + \varepsilon_{it} \quad (2)$$

$$Y_{it}^{nat} = \beta_0 + \gamma_f^{nat} + \beta_1 X_{it} + \varepsilon_{it} \quad (3)$$

Where equation (2) includes all graduates with positive in-state earnings, and equation (3) includes all graduates with positive national earnings. The key coefficients of this regression are γ_f^{in} and γ_f^{nat}, which are the major fixed effects for in-state and national earnings regressions, and measure the field-specific returns. We omit psychology, as it is the largest field, and also relatively generic, so all the coefficients are in reference to the earnings of a psychology degree. To summarize the results from these regressions, we include Figure 10, which shows the difference between the national and in-state fixed effects for the same majors that we display in Figure 5. Panel (a) shows results that include all the quarterly earnings, while panel (b) shows results from separate regressions for 1, 5 and 10 years after graduation. Consistent with the results from Figure 5, Social Sciences and Business returns are understated when using in-state earnings by almost 7 percentage points relative to the return for Psychology. Additionally, these biases are larger compared to Psychology in the first year, when very few psychology majors move out of the state. In this first year the in-state returns understate national returns by over 10 percentage points. In the year 10 results, the biases are smaller, likely because many Psychology majors leave the state, particularly at the high end of the earnings distribution (Figure 5, panel b). The implication is that researchers estimating earnings differences across programs will need to confront the likely differential migration between majors which may under- or over-state earnings differences across fields and to an extent that changes with the time horizon.

V. Empirical Tests, Bounding, and Other Corrections

A. Empirical Tests and Supplemental Data

Prior researchers have proposed a few empirical tests for the presence of migration bias and also brought in supplemental data. Here we describe these approaches. Subsequent work will evaluate the ability of these approaches to distinguish settings with minimal from large bias.

1. Is treatment associated with having non-missing earnings?

The most common test is whether treatment is associated with the likelihood of having non-zero in-state earnings. While a failure of this test could suggest bias, no difference does not ensure that there is no selection bias. The treatment and control groups may simply be experiencing differential (but equally-sized) selection. Zimmerman (2014) tests whether treatment is associated with "In LF sample" and finds a 2 pp decrease in likelihood of positive earnings, interpreted as small. However, if inflated by the first stage, this difference is 8 pp reduction in sample inclusion associated with admission. Denning, Marx, and Turner (2018) test whether treatment affects the probability of having either in-state earnings or enrollment, finding a small positive association (< 1 pp for early years, > 1 for year seven). These estimates would be larger if appropriately scaled by the first stage. Having an automatic zero EFC and more financial aid (treatment) makes students more likely to be observed in-state. Note that this does not fully capture the extent of moving because a person whose move straddles two program years will have zero earnings in some but not all quarters. It also combines the effect on likelihood of being in-state college, working (vs. not working), and working in-state (vs. working out-of-state). Ost, Pan, and Webber (2018), finding no association between treatment and attrition, nicely sum up the limitation of Lee (2007) bounds to this scenario: "given that there is no evidence of differential attrition to begin with, it is no surprise that our results are robust to [the Lee (2007)] bounding exercise." Finally, Altonji and Zimmerman (2019) reports difference by field of study (relative to education, baseline 0.128 missing) ranging from -1 to +25 percentage points.

Our findings in Section IV.B also illustrate the issues with using differential attrition as the main test for the presence of bias, given that UT-Austin and CU-Boulder have similar differences in attrition as other schools in the state, but very different levels of earnings bias.

2. Balance Tests for Full vs. Restricted Samples

Some authors demonstrated the balance of covariates between treatment and control groups for the selected sample with non-zero earnings. Ost, Pan, and Webber (2018) and Zimmerman (2014) perform such a test, finding that covariates are still balanced in their RD setting. Andrews, Li, and Lovenhiem (2016) present means of covariates for treatment and control groups separately for those included and excluded in the analysis due to lack of earnings observations. While they do not present formal tests, there does not appear to be any differential attrition between treatment and control groups based on these covariates. However, neither of these rules out the possibility of differential attrition due to unobserved factors, most importantly latent earnings offers. We present such a test for our sample in Table 5. We examine whether the extent of covariate balance between flagship and non-flagship graduates differs between the full and in-state earnings samples. Though treatment is not balanced on covariates (as expected, given our lack of quasi-experimental variation), the extent of balance does not appear to differ between the full and restricted samples. This suggests this test does not differentiate settings with minimal vs. substantial bias.

3. *Supplemental data: How much migration is there? Is it associated with treatment or outcomes?*

While researchers rarely have access to migration data specifically pertaining to their sample, supplemental data such as the American Community Survey (ACS) or the Current Population Survey (CPS) can be informative. Using the ACS, Ost, Pan, and Webber (2018) estimate that non-earnings among individuals with at least some college is half attributable to leaving the state in the last year (56%), a third due to true non-employment (32%), and the remainder due to self-employment (8%) and federal government employment (4%). From the IPUMS-CPS, Denning, Marx, and Turner (2018) estimate an annual interstate migration rate of 3.2 percent for young adults with some college from Texas between 2010 and 2016. Compounding these annual rates over a decade could result in substantial migration from the state, though this data is not able to determine differential rates between treatment and control groups. Andrews, Li, and Lovenheim (2016) provide the best illustration of this approach. In the 2000 Census, they identify recent college graduates who lived in Texas five years earlier (when they were aged 17-21). They use living in the Austin or College Station MSAs (vs. rest of Texas) as proxies for having graduated from the flagships UT-Austin and Texas A&M, respectively, which correspond to their treatments of interest. They then document the log earnings distribution for these workers separately by Texas MSA and whether the workers are in or out of Texas. For all three MSA groups (Austin, College Station, rest of Texas), the in- and out-of-state earnings distributions are similar, suggesting that higher earners are not more likely to move out-of-state, whether from a flagship or not. While quite encouraging, any error in the measurement of treatment status will tend to attenuate differences. Furthermore, this approach is simply not available for other treatments. For example, Boulder is not separately identifiable in the Census or ACS.

B. Is Bounding Appropriate? How Does It Perform?

Lee (2009) proposed a bounding approach to estimate treatment effects in the presence of sample attrition. He developed the approach to estimate treatment effects on wage rates (rather than total earnings) in the presence of non-random employment: wages are only available for people who work, so conditioning on working introduces sample selection bias. The idea is to exclude individuals from the group that experiences less attrition so that treatment and control groups are comparable on the remaining distribution. Subsequent work has applied the approach to more general settings where sample attrition is correlated with treatment. This would seem a natural approach to dealing with attrition in our setting, as treatment is highly correlated with the likelihood of observing non-zero earnings, because treatment affects both employment and migration. The key assumption to this approach is monotonicity: treatment must only affect attrition in one direction. In Lee's case, the treatment effect on the employment probability is assumed to have the same sign for all individuals. This assumption rules out that treatment may increase employment for some individuals, while reducing it for others. In our case, there are good reasons to think that the monotonicity assumption would be violated since individuals can attrit on two margins: employment and out-of-state migration. Monotonicity would be violated if treatment increased employment for some individuals and increased out-of-state migration for others, which seems likely.

Nonetheless, Table 6 implements this bounding approach in our setting. Note that the resulting coefficients are not directly comparable to our results in Table 3 as they do not include any controls, including for state, graduating cohort, year, or quarter. Given the large difference in match rates between flagship and non-flagships (combining both migration and non-employment), this procedure trims a large share of the sample (20% of the non-flagship group).

As a consequence, the trimming produces very wide bounds, ranging from -0.20 to + 0.37 log points combining all time periods. The range still includes zero at ten years out.

Lee (2007, p. 1096) offers an additional test of the monotonicity assumption explicitly: "If it was found that for some values of X, the treatment caused wages to be observed, while for other values of X, the treatment was found to cause wages to be missing, then the monotonicity assumption must not hold". For instance, if high ability students experience a reduction in sample inclusion if they get treatment while low ability students experience an increase in sample inclusion if they get treatment, then this suggests a failure of monotonicty assumption. While this does not indicate the presence of migration bias, a failure to reject the monotonicity assumption lends support to the Lee (2007) bound approach. Prior work has not applied this test to assessing the potential for migration bias.

C. Other Approaches and Corrections

Various approaches to interpolating or imputing earnings for non-matched records appear to increase bias. Prior work has used inverse probability weighting to correct for attrition in survey work, which relies attrition to be based on observables. Grogger (2013) proposes using runs of zeros at the end of the sample period in administrative data to construct bounds on interstate mobility, which then can be used to bound treatment effects. Finally, it may be possible to parameterize the bias as a function of factors observable to researchers, such as jurisdiction size, unemployment rate, and participant age so that analysts can determine whether their setting is likely to be one with high bias. The evaluation of these other approaches is ongoing.

VI. Monte Carlo Evidence

To examine the extent of bias under a more general set of conditions than our empirical example, we develop and simulate a simple model of earnings, work, and migration in the presence of some treatment.

A. Simulation model setup

Each person is characterized by six random variables:

- Treatment status (T) is randomly assigned, allowing us to abstract from bias arising from non-random selection into treatment.
- Draw from an ability distribution $A \sim N(0, \sigma_{ability})$
- Draw from in-state earnings offer distribution: $\tilde{y}_{in} = \beta_0 + A + \beta_{in}T + \varepsilon_{in}$ where $\varepsilon_{in} \sim N(0, \sigma_{in})$
- Draw from out-of-state earnings offer distribution $\tilde{y}_{out} = \beta_0 + A + \beta_{out}T + \varepsilon_{out}$ where $\varepsilon_{out} \sim N(0, \sigma_{out})$
- Reservation wage $r \sim U(0, r_{max})$ which is the same for in-state and out-of-state jobs
- Moving cost: $c \sim U(0, c_{max})$, which is uncorrelated with reservation wages and job offers

Note that in-state and out-of-state earnings offers are correlated both through the inclusion of ability A in both distributions and because treatment influences the means of both distributions (by β_{in} and β_{out}, respectively). This model assumes treatment effect homogeneity on earnings offers, though there will be heterogeneity in actual earnings effects depending on an individual's reservation wage and moving cost. Individuals with a high reservation wage will have a lower treatment effect because they will be less likely to move from non-employment.

Labor force participation decisions are made (separately for in-state and out-of-state earnings offers) by comparing offered earnings to the reservation level. Thus accepted earnings in each labor market are truncated:

$$y_{in} = \tilde{y}_{in} * 1\{\tilde{y}_{in} > r\} \text{ and } y_{out} = \tilde{y}_{out} * 1\{\tilde{y}_{out} > r\}$$

Mobility decisions are made by comparing the difference in accepted offers between labor markets to moving costs:

$$Move = 1\{y_{out} - y_{in} > c\}$$

Finally, actual earnings is given by

$$y_{actual} = y_{in}(1 - Move) + y_{out}Move$$

The problem arises in that y_{actual} is not observed by the researcher, but rather earnings are observed as zero if the worker moves out of state:

$$y_{observed} = y_{in}(1 - Move) + 0 * Move$$

To illustrate the bias that arises in such a model, we simulate 100,000 draws with the following parameters: β_0=\$8000, $\sigma_{ability}$=\$2000, σ_{in}=\$1000, σ_{out}=\$1000, r_{max}=\$8000, c_{max}=\$3000. We set β_{in}=\$2000, which corresponds to a 25% treatment effect on the mean of the in-state earnings distribution. We'll see that results are particularly sensitive to the relative treatment effects on out-of-state and in-state earnings, so we present results where β_{out} equals different multiples of β_{in}. We also examine simulations in which migration is exogenous.

B. Simulation results

In Table 7 we report results of regressions using this simulated data. Panel A depicts our base model, where treatment differentially increases out-of-state earnings ($\beta_{out} > \beta_{in}$) and migration is endogenous in the sense of responding to earnings differentials across areas. The most naïve comparison – simply comparing observed earnings between treatment and control group – is very negatively biased. In fact, the point estimate is close to zero when in fact the true effect of treatment is \$2874.[19] This is because out-of-state workers have higher earnings as a consequence of the treatment, but are coded as having no earnings. Many researchers restrict the analysis sample to workers with non-zero in-state earnings. Doing so lowers, but does not eliminate the migration bias (second row). Furthermore, it should be noted that this restriction changes the estimand to the effect of treatment on earnings conditional on (non-random) participation. Ignoring the extensive earnings margin will understate the total earnings (and welfare) effect of an intervention. With this caveat, we continue to focus on this estimand. Estimates of effects on log earnings (restricting to individuals with positive earnings) will also be biased downwards, particularly at the high end of the distribution. The bottom of the table describe migration patterns for the sample. Moving is highly correlated with both treatment and earnings: treatment is associated with a 20 percentage point increase in likelihood of moving and \$1000 more in actual earnings is associated with a 4 percentage point increase in likelihood of moving. Individuals that move have earnings that are \$1878 higher than those that don't. This suggests

[19] Note that the true treatment effect is a weighted average of the treatment effects on the in- and out-of-state earnings offer distributions (2000 and 3000, respectively, in Panel A) combined with any effects on migration and labor supply.

two conditions for the presence of bias: migration is related to earnings and treatment is related to migration.

Panel B shows results from a simulation where treatment does not differentially affect in- and out-of-state earnings offers ($\beta_{out} = \beta_{in}$). The naïve model is still biased downwards slightly because out-of-state workers have higher earnings (moving costs must be overcome) and these higher earnings are erroneously set to zero. However, there is no bias in any other specifications. While migration is still endogenous and related to actual earnings (higher earning individuals are more likely to move), treatment is now unrelated to moving. Non-random migration will still affect estimates of the overall earnings distribution, but treatment effect estimates will not be subject to bias.

Finally, Panel C shows results from a simulation where treatment differentially affects in- and out-of-state earnings offers ($\beta_{out} > \beta_{in}$), but migration is exogenous (set at 25%). That is, individuals are randomly assigned to move regardless of their earnings offers, moving costs, or treatment status. While this feature reduces the bias relative to the base case, it does not eliminate it. Indeed even with exogenous migration there is still an association between migration and actual earnings because treatment effects out-of-state earnings. Thus any effect of treatment on earnings that only occurs on out-of-state earnings is lost when you only use in-state students.

To better understand the mobility patterns that give rise to these results, Figure 11 depicts the simulated share of individuals that move out of state by rank on the true earnings distribution, separately for the treatment and non-treated groups. The panels correspond to those in Table 7. The base model (Panel A) shows that migration is clearly related to actual earnings, with higher earners more likely to have moved out-of-state. Importantly, individuals in the treatment group have a steeper gradient with earnings than the control group. Thus the observed earnings distribution of treated individuals will systematically be missing high earners relative to the control group, causing bias in treatment effect estimates. This pattern roughly corresponds to the flagship treatment example discussed previously. Interesting, this pattern remains even if migration is exogenous (Panel C). However, when treatment has no differential effect (Panels B and D), the migration-earnings relationship is similar for treatment and control groups, eliminating bias. Even a modest relationship between earnings and migration, as seen in Panel B, will not necessarily cause bias as long as it is similar in the treatment and control groups.

In Figures 12 through 14 we examine how the mobility patterns and resulting bias change as we shift the differential earnings effect of treatment (ratio β_{out}/β_{in}). Figure 12 plots the overall share moving and how moving correlates with treatment and actual earnings, separately for endogenous and exogenous migration. First consider endogenous migration. The probability of moving increases as treatment has a larger effect on out-of-state earnings, as does the correlation between moving, treatment, and actual earnings. Figure 13 plots the resulting bias for the five earnings outcomes. Using the full sample (but assigning out-of-state earnings as zero) results in considerable bias, even when there is no differential ($\beta_{out} = \beta_{in}$). Conditioning the sample on workers with non-zero earnings (which effectively drops movers) reduces bias. Bias increases with increases in the differential effects on in- and out-of-state earnings, though it is asymmetric. When treatment favors in-state earnings ($\frac{\beta_{out}}{\beta_{in}} < 1$) bias is positive, low, and stable. Bias becomes increasingly negative as $\frac{\beta_{out}}{\beta_{in}} > 1$, both in earnings level and percentiles. This asymmetry arises from the asymmetry of the moving decision; moving costs are always positive so the likelihood that the treatment group is truncated is lower when moving is less and less desirable ($\frac{\beta_{out}}{\beta_{in}} < 1$).

Now consider exogenous migration (Panel B of Figure 12). Though the migration rate does not depend on $\frac{\beta_{out}}{\beta_{in}}$ by construction, its correlation with actual earnings does. Figure 14 depicts the bias with exogenous migration. Bias is positive when $\frac{\beta_{out}}{\beta_{in}} < 1$ and negative when $\frac{\beta_{out}}{\beta_{in}} > 1$, approximately linear in the ratio and symmetric. Exogenous migration breaks the dependence of migration on the relative treatment effect that underlies the asymmetry when moving is endogenous.

C. Bounding Approaches

We examine the performance of some alternative bounding approaches with our Monto Carlo simulations and report results in Table 8. In the first, we use the full sample of individuals, including those with no matched earnings (due to either non-employment or migration). We substitute zero earnings for the actual earnings of individuals in the top or bottom D% of the non-zero earnings distribution of the control group. D is the difference in match rates between the two groups as a proportion of the control group match rate. In our simulation D equals 17%, with the control group more likely to match. This generates an upper (lower) bound of the true treatment effect under the extreme assumption that all of difference in match probabilities comes from untreated individuals with the highest (lowest) earnings who would have otherwise left the state if they were treated, assuming no effect of treatment on employment.[20] Panel B reports these results. While an improvement over the naïve regression, the bounds [$796, $1568] nonetheless do not contain the actual treatment effect ($2874). The upper bound fails to capture any earnings improvement operating via increased employment.

In Panel C, we specifically implement Lee's (2009) approach by restricting our analysis to records with non-zero (positive) observed earnings. We then omit the top (bottom) 17% of the control group observed earnings distribution when calculating the treatment-control outcome difference. We do this for mean earnings levels, log earnings, and for moments of the log earnings distribution. The constructed bounds do contain the true parameter (as well as the biased estimated parameter) in all cases, though the truth is typically closer to the upper than lower bound and the bounds are wide.

In Panels D and E we implement sharper Lee (2009) bounds by introducing a baseline (pre-treatment) covariate. The process essentially involves computing bounds separately for twenty groups defined by individual's latent ability (Panel D) or moving cost (Panel E), then computing a weighted average of these group-specific estimates. Latent ability, which is much more highly correlated with earnings than moving costs, tightens the bounds considerably. However, the upper bound [$2,390] nearly omits the true effect [$2,315]. Finally, in Panel F we implement bounds that are robust to a failure of the monotonicity assumption, as suggested by Zhang and Rubin (2003): we trim both the treatment and control groups by their rates of missing in-state employment. Given the high rates of non-employment and out-migration, this approach yields bounds that are completely uninformative.

D. Lessons from the simulations

[20] This is the approach taken by Denning, Marx, and Turner (2018), though the rate of differential attrition in their setting is much lower than our simulations.

We take four lessons from this simulation analysis. First, the ratio $\frac{\beta_{out}}{\beta_{in}}$ is a key determinant of bias. Bias is zero when treatment has a similar effect on in-state and out-of-state earnings ($\beta_{out} = \beta_{in}$), as this is what induces differential migration by earnings in the treatment and control groups. Second, there can be bias even if migration is completely exogenous. Again, differential treatment effects for in-state and out-of-state earnings will truncate the observed earnings distribution of the treatment group more than the control group even if migration is exogenous. Third, bias is reduced when the sample is conditioned on having positive observed earnings. This drops both movers and in-state non-participants, so it does change the target of estimation to a parameter that does not fully capture the consequences of the treatment under study. Finally, a test of the presence of bias is whether the relationship between migration and earnings differs between the treatment and control groups. Interestingly, having migration unrelated to treatment (the exogenous mobility case in Panel C) is not sufficient to rule out bias. Nor is a relationship between moving and actual earnings necessarily evidence of bias (the endogenous mobility with $\beta_{out} = \beta_{in}$ in Panel B).

VII. Conclusion

Many papers seek to estimate the effect of an educational treatment on earnings. However, many of those papers use administrative earnings records from a single state, and are thus restricted to measure earnings outcomes using individuals that remain and work in-state. While most authors have recognized the potential for bias, prior work has not had access to data that would permit them to directly test the extent of bias. This study takes advantage of a unique match between postsecondary records from two states with administrative records nationally, permitting us to quantify the extent of bias due to out-of-state migration.

Using the effect of flagship enrollment as an example, we find considerable bias from the use of in-state earnings records exclusively. We conclude that the flagship effect is actually 26% higher than that suggested by in-state administrative earnings records, at least in our context. Migration bias also confounds estimates of earnings differences across majors. Simulations show that this bias can arise *even if migration itself is random*, as long as the distribution of earnings is different for treated and non-treated individuals. Additionally, we show that this attrition bias appears to affect the right tail of the earnings distribution, and that the issue worsens over time.

We also evaluate the performance of various strategies (e.g. Lee, 2008) commonly used to deal with attrition bias, though such bounding exercises are uninformative in this setting. As the use of administrative data continues to proliferate, a better understanding of the bias resulting from inter-jurisdiction migration and how to address it will be invaluable. Alternative approaches to test for and correct for migration-related bias are still being evaluated.

Our analysis has two implications for states. First, states do not retain many of their highest-paid university graduates, which is a goal of many state merit-aid programs. An inability to follow students that leave the state is a big barrier to evaluating the ability of such programs to retain talent. Second, earnings estimates published by state higher education boards and made available to students will understate earnings differences between programs (institutions and fields) due to systematic differences in rates of out-migration. Published earnings records for smaller states with high rates of out-migration will be particularly misleading.

References

Abowd, J. M., Stephens, B. E., Vilhuber, L., Andersson, F., McKinney, K. L., Roemer, M., & Woodcock, S. (2009). The LEHD infrastructure files and the creation of the Quarterly Workforce Indicators. In *Producer dynamics: New evidence from micro data* (pp. 149-230). University of Chicago Press.

Altonji, J. and S. Zimmerman 2018. The Costs of and Net Returns to College Major. In C. Hoxby and K. Stange eds., *Productivity in Higher Education*, forthcoming, Chicago, Illinois: University of Chicago Press.

Anelli, M, 2018. The Labor Market Determinants of the Payoffs to University Field of Study. Unpublished working paper. February 15, 2018.

Andrews, R, J. Li, and M. Lovenheim, 2016. "Quantile Treatment Effects of College Quality on Earnings. *Journal of Human Resources*, 51(1): 201-238.

Andrews, R. and K. Stange, 2019. Price Deregulation and Equality of Opportunity in Higher Education: Evidence from Tuition Deregulation in Texas. *American Economic Journal: Economic Policy*, forthcoming.

Artmann, Elisabeth, Nadine Ketel, Hessel Oosterbeek, Bas van der Klaauw, 2018. Field of Study and Family Outcomes. IZA DP No. 11658

Bahr, P. R., Dynarski, S., Jacob, B., Kreisman, D., Sosa, A. Wiederspan, M, 2015. Labor market returns to community college awards: Evidence from Michigan. CAPSEE working paper

Bakkes Holzer Valez (2015) Backes, B. , Holzer, H. J. , & Velez, E. D. 2015. Is it worth it? Postsecondary education and labor market outcomes for the disadvantaged. *IZA Journal of Labor Policy*, 4 (1).

Barnow, Burt and David Greenberg, 2015. "Do Estimated Impacts on Earnings Depend on the Source of the Data Used to Measure Them? Evidence From Previous Social Experiments." *Evaluation Review*. Vol. 39(2) 179-228.

Behaghel, Luc, Bruno Crépon, Marc Gurgand, and Thomas Le Barbanchon, 2015. Please Call Again: Correcting Non-Response Bias in Treatment Effect Models. *The Review of Economics and Statistics*, December 2015, 97(5): 1070–1080

Behaghel, L., B. Crépon, M. Gurgand, and T. Le Barbanchon, 2012. "Please Call Again: Correcting Non-Response Bias in Treatment Effect Models," IZA discussion paper 6751.

Belfield, C, J. Britton, F. Buscha, L. Dearden, M. Dickson, L. Van der Erve, L. Sibieta, A. Vignoles, I. Walker, and Y. Zhu, 2018. "The relative labour market returns to different degrees." Institute for Fiscal Studies Research Report, June 2018.

Bollinger, C., B. Hirsch, C. Hokayem, J. Ziliak, 2018. Trouble in the Tails? What We Know about Earnings Nonresponse Thirty Years after Lillard, Smith, and Welch. *Journal of Political Economy*. Forthcoming.

Britton, Jack, N. Shephard, and A. Vignoles, 2018. "A comparison of sample survey measures of English graduates with administrative data." *Journal of the Royal Statistical Society*, A (2019) 182 (Part 3): 1-30.

Canaan, Serena, and Pierre Mouganie. 2018. Returns to Education Quality for Low-Skilled Students: Evidence from a Discontinuity. *Journal of Labor Economics*. 36(2).

Card, D., R. Chetty, M. Feldstein, and E. Saez, 2010. Expanding Access to Administrative Data for Research in the United States. *National Science Foundation White Paper* 10-069 September 2010.

Carruthers, C. and T. Sanford, 2018. Way station or launching pad? Unpacking the returns to adult technical education. *Journal of Public Economics*. 165: 146-159

Chetty, Raj. 2012. "Time Trends in the Use of Administrative Data for Empirical Research" presentation at NBER Summer Institute, July 2012.

Chyn, Eric. 2018. "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children." *American Economic Review* 2018, 108(10): 3028–3056

Council of Economic Advisors, Executive Office of the President of the United States, 2017. "Using Federal Data to Measure and Improve the Performance of U.S. Institutions of Higher Education." Updated January 2017.

Cunha, J. M. and T. Miller, 2014. Measuring value-added in higher education: possibilities and limitations in the use of administrative data. *Economics of Education Review* 42, 64–77.

Denning, J., B. Marx, and L. Turner, 2018. ProPelled: The Effects of Grants on Graduation, Earnings, and Welfare, *American Economic Journal: Applied Economics*, forthcoming.

Doyle, Joseph. 2013. "Causal effects of foster care: An instrumental-variables approach" *Children and Youth Services Review* 35 (2013) 1143–1151

Dyke, A. , Heinrich, C. J. , Mueser, P. R. , Troske, K. R. , & Jeon, K.-S. 2006. The effects of welfare-to-work program activities on labor market outcomes. *Journal of Labor Economics*, 24 (3), 567–607

Engbom N. and C. Moser, 2017. Returns to Education Through Access to Higher-Paying Firms: Evidence from US Matched Employer-Employee Data. Unpublished working paper April 8, 2017.

Figlio, David, Krzysztof Karbownik, and K. Salvenes, 2017. “The Promise of Administrative Data in Education Research” *Education Finance and Policy*, Presidential Essay. 2017.

Fitzpatrick, Maria D. and A. Damon Jones. 2016. “Higher Education, Merit-Based Scholarships and PostBaccalaureate Migration. *Economics of Education Review*. 54: 155-172.

Franklin, Rachel S. 2013. "Domestic Migration Across Regions, Divisions, and States: 1995 to 2000" Census 2000 Special Reports.

Grogger, Jeffrey. 2013 “Bounding the Effects of Social Experiments: Accounting for Attrition in Administrative Data.” *Evaluation Review* 36(6) 449-474.

Hastings, J.S., C.A. Neilson, and S.D. Zimmerman, 2013. “Are Some Degrees Worth More Than Others? Evidence from College Admissions Cutoffs in Chile,” NBER Working Paper, 2013, 19241.

Hoekstra, Mark. 2009. “The Effect of Attending the State Flagship University on Earnings: A Discontinuity- Based Approach.” *Review of Economics and Statistics* 91(4):717–24.

Jepsen, Christopher, Kenneth Troske, and Paul Coomes. 2014. “The Labor- Market Returns to Community College Degrees, Diplomas, and Certifi cates.” *Journal of Labor Economics* 32(1):95–121.

Kirkebøen, Lars, Edwin Leuven, and Magne Mogstad, 2016. “Field of Study, Earnings, and Self-Selection,” *The Quarterly Journal of Economics*, 2016, p. qjw019.

Lachowska, Marta, Alexandre Mas, Stephen A. Woodbury, 2018. “Sources of Displaced Workers' Long-Term Earnings Losses” NBER Working Paper No. 24217 Issued January 2018.

Lee, D., 2009. “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects,” *Review of Economic Studies* 76: 1071–1102.

Lillard, Lee, James P. Smith, and Finis Welch. 1986. “What Do We Really Know about Wages? The Importance of Nonreporting and Census Imputation,” *Journal of Political Economy* 94 (June 1986): 489-506.

Liu, V. Y. T., Belfield, C. R. , & Trimble, M. J. 2015. The medium-term labor market returns to community college awards: Evidence from North Carolina. *Economics of Education Review*, 44 , 42–55 .

Malamud, O. and A. Wozniak, 2012. The Impact of College on Migration Evidence from the Vietnam Generation. *Journal of Human Resources.* 47(4): 913-950.

Minaya, V. and Judith Scott-Clayton, 2018. Labor Market Outcomes and Postsecondary Accountability: Are Imperfect Metrics Better than None? In C. Hoxby and K. Stange eds., *Productivity in Higher Education*, forthcoming, Chicago, Illinois: University of Chicago Press.

Mueller-Smith, Michael, 2018. "The Criminal and Labor Market Impacts of Incarceration." *Working paper. University of Michigan.*

Organization for Economic Cooperation and Development (OECD), 2015. *Connecting with Emigrants - A Global Profile of Diasporas 2015*. Table 4.2 total emigration rates and emigration rates of the highly skilled, by country of origin, 2010/2011.

Ost, Ben, Weixiang Pan, and Douglas Webber. 2018. The Returns to College Persistence for Marginal Students: Regression Discontinuity Evidence from University Dismissal Policies. *Journal of Labor Economics*. 36 (3): 779-805.

Penner, Andrew M. & Kenneth A. Dodge, 2019. "Using Administrative Data for Social Science and Policy" *RSF: The Russell Sage Foundation Journal of the Social Sciences* 5(3): 103–27. DOI: 10.7758/RSF.2019.5.3.06.

Ransom, T., 2016. "Selective Migration, Occupational Choice, and the Wage Returns to College Majors." Unpublished working paper.

Scott-Clayton, Judith and Q. Wen, 2017. "Estimating Returns to College Attainment: Comparing Survey and State Administrative Data Based Estimates." CAPSEE Working Paper. January 2017.

Stevens, A., M. Kurleander, and M. Groz, 2018. Career Technical Education and Labor Market Outcomes: Evidence from California Community Colleges. *Journal of Human Resources*. Forthcoming.

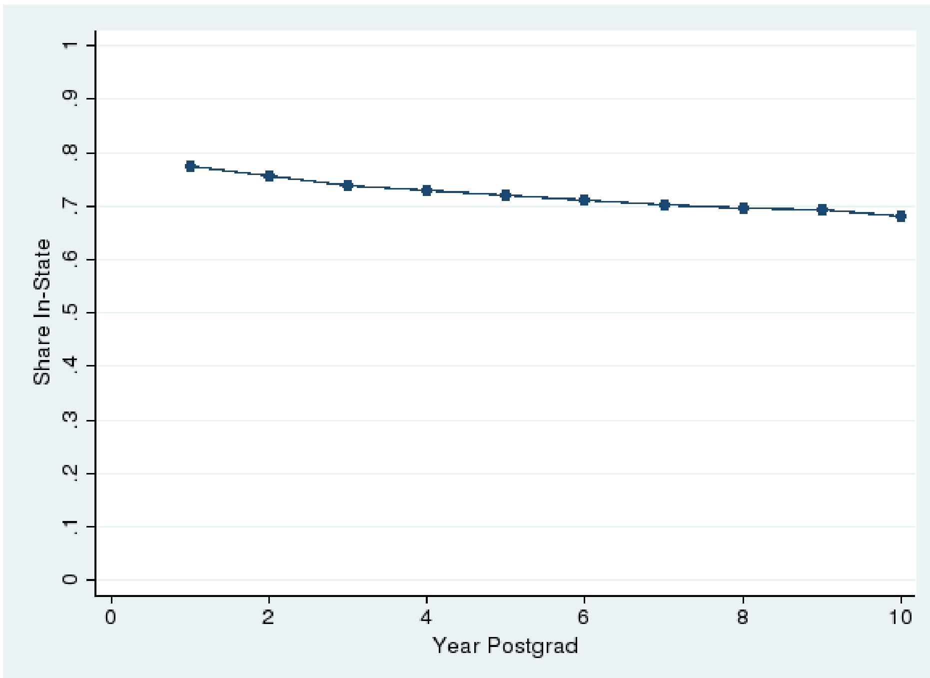
Turner, L. 2015. The returns to higher education for marginal students: Evidence from Colorado Welfare recipients. *Economics of Education Review* 51 (2016) 169–184

Walker, I. and Zhu, Y. 2011. Differences by degree: evidence of the net financial rates of return to undergraduate study for England and Wales. *Economics of Education Review*, 30, 1177–1186.

Walker, I. and Y. Zhu, 2017. University selectivity and the graduate wage premium: evidence from the UK. IZA Discussion Paper 10536.

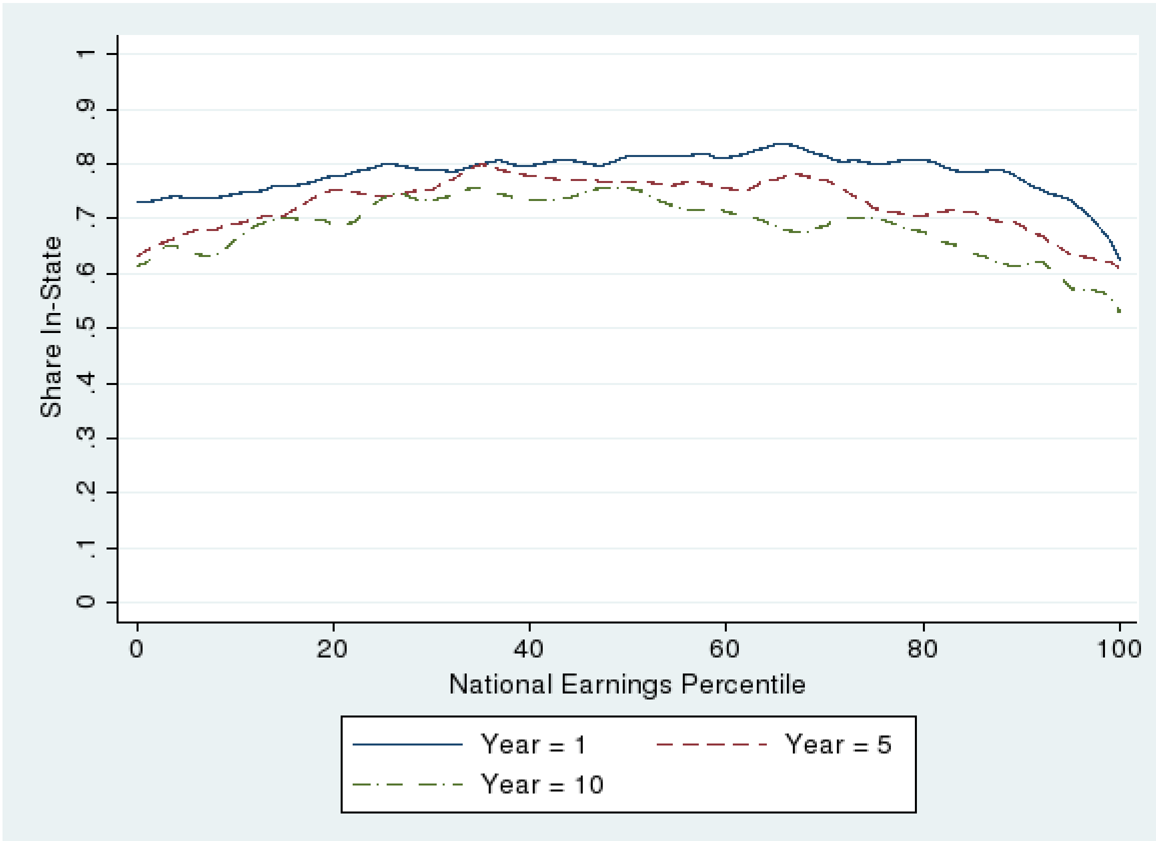
Zimmerman, Seth. 2014. "The Returns to College Admissions for Academically Marginal Students." *Journal of Labor Economics* 32(4):711–54.

Figure 1: Share of Employed UT System Graduates In-State



Note: Author's calculations using merged data from University of Texas System and earnings data from the LEHD. Sample is 10% of UT graduates who have at least three quarters of non-zero earnings and with at least $10,000 of earnings in the calendar year nationally (in-state or out-of-state).

Figure 2: Share In-State by Percentile in National Earnings Distribution and Year Post-Grad

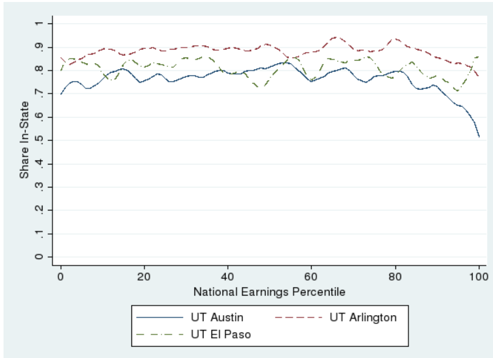


Note: Author's calculations using merged data from education records and earnings data from the LEHD. Sample is 10% of graduates from the University of Texas or public universities in Colorado who have at least three quarters of non-zero earnings and with at least $10,000 of earnings in the calendar year nationally (in-state or out-of-state). National earnings percentile is defined relative to other graduates in the sample observed in years 1, 5, and 10, respectively.

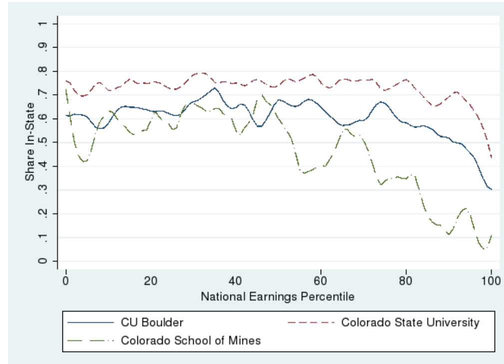
Figure 3: Share In-State by Percentile in Lifetime Earnings Distribution and Year Post-Grad

Note: Author's calculations using merged data from UT and Colorado institutions and earnings data from the LEHD. Sample is 10% of graduates who have at least three quarters of non-zero earnings and with at least $10,000 of earnings in at least 10 years. National earnings percentile is defined relative to other graduates' lifetime earnings in the sample.

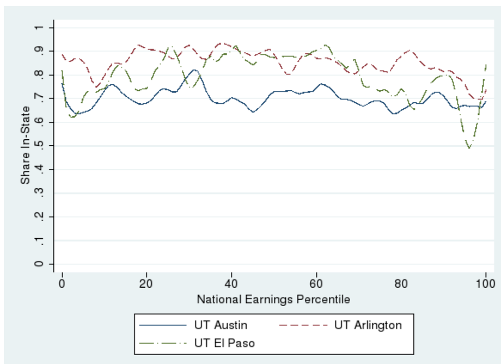
Figure 4: Share In-State by Percentile in the Earnings Distribution, by Institution



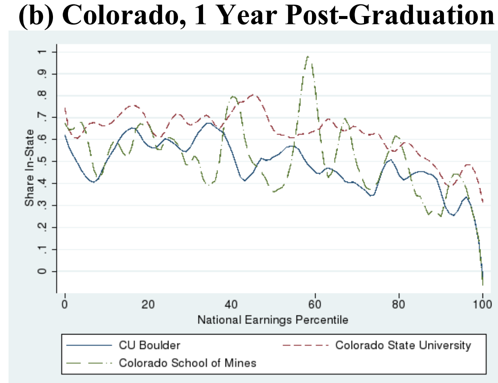
(a) Texas, 1 Year Post-Graduation



(b) Colorado, 1 Year Post-Graduation



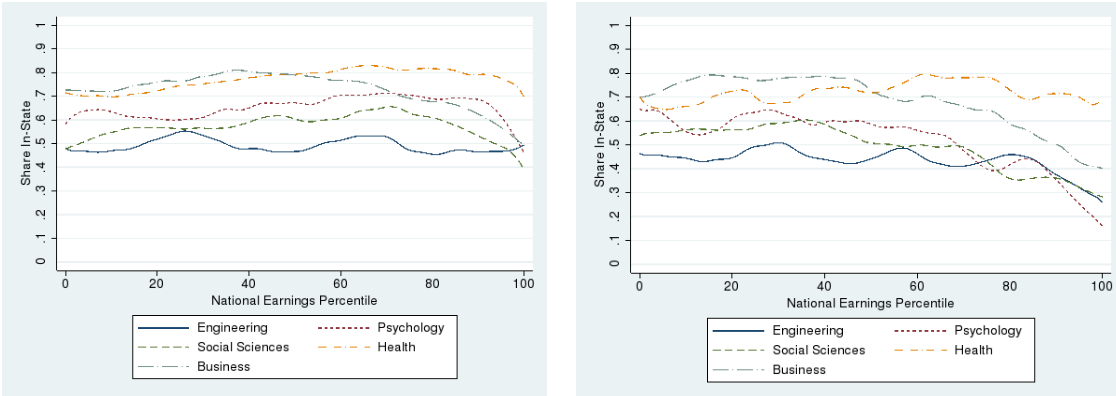
(c) Texas, 10 Years Post-Graduation



(d) Colorado, 10 Years Post-Graduation

Note: Author's calculations using merged data from education records and earnings data from the LEHD. Sample is 10% of graduates from the University of Texas or public universities in Colorado who have at least three quarters of non-zero earnings and with at least $10,000 of earnings in the calendar year nationally (in-state or out-of-state). National earnings percentile is defined relative to other graduates in the sample from the same institution.

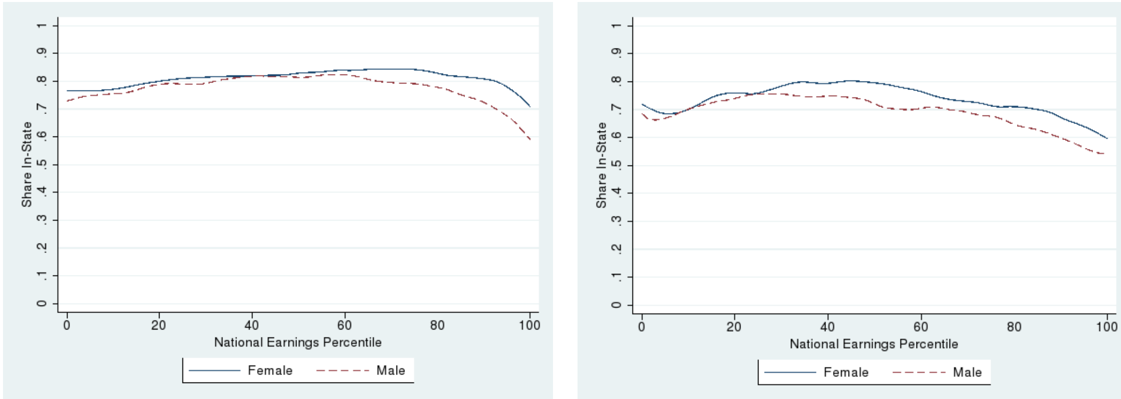
Figure 5: Share In-State by Percentile in the Earnings Distribution, by Major



(a) 1 year post-graduation **(b) 10 years post-graduation**

Note: Author's calculations using merged data from education records and earnings data from the LEHD. Sample is 10% of graduates from the University of Texas or public universities in Colorado who have at least three quarters of non-zero earnings and with at least $10,000 of earnings in the calendar year nationally (in-state or out-of-state). National earnings percentile is defined relative to other graduates in the sample in the same major.

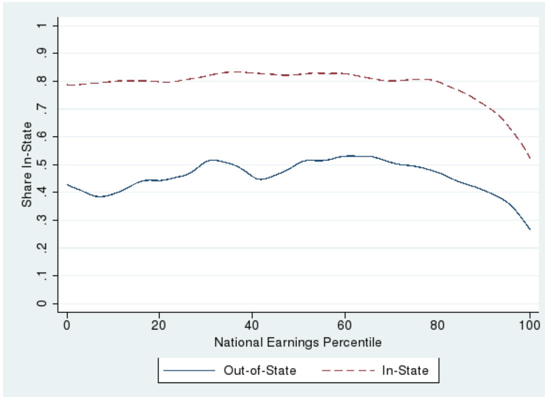
Figure 6: Share In-State by Percentile in the Earnings Distribution, by Gender



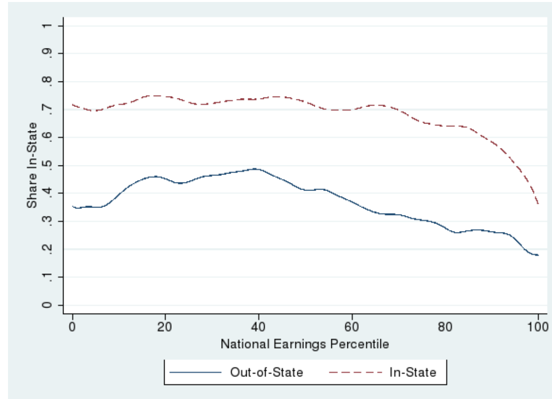
(a) 1 year post-graduation **(b) 10 years post-graduation**

Note: Author's calculations using merged data from education records and earnings data from the LEHD. Sample is 10% of graduates from the University of Texas or public universities in Colorado who have at least three quarters of non-zero earnings and with at least $10,000 of earnings in the calendar year nationally (in-state or out-of-state). National earnings percentile is defined relative to other graduates in the sample with the same gender.

Figure 7: Share In-State by Percentile in the Earnings Distribution, by Major



(a) 1 year post-graduation

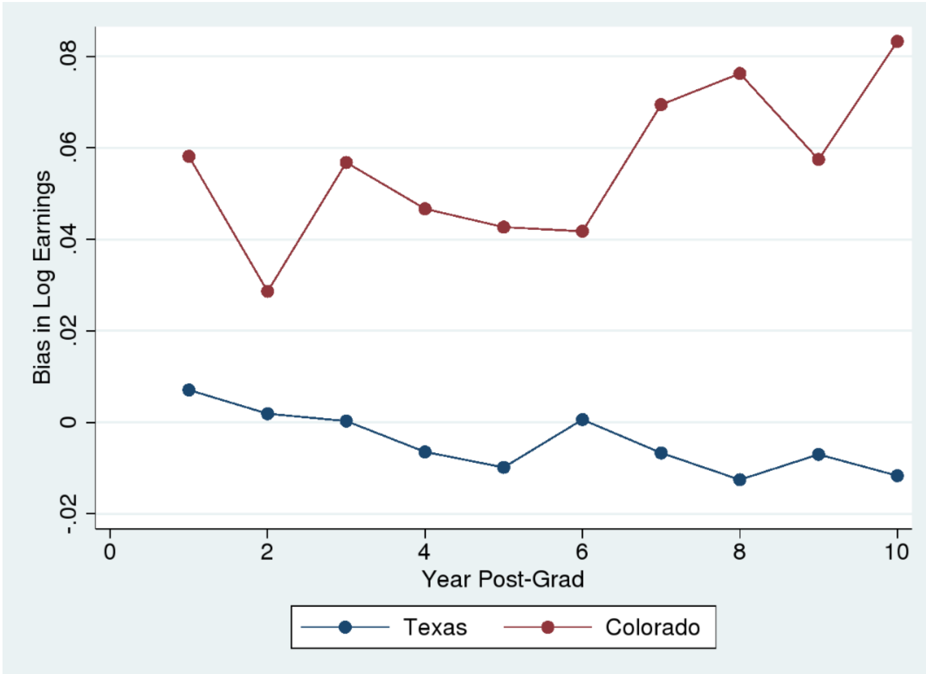


(b) 10 years post-graduation

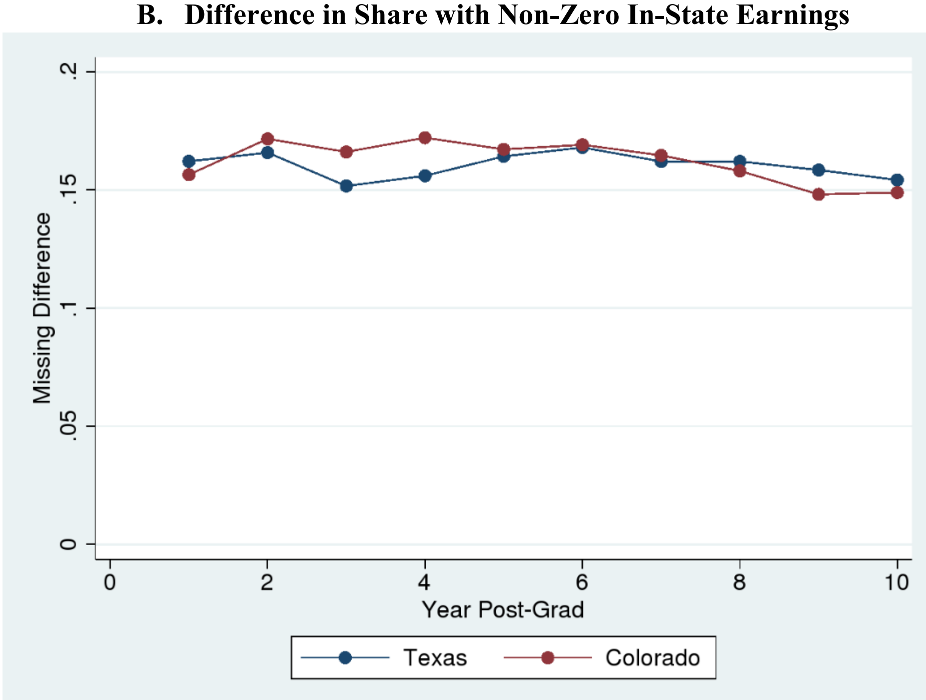
Note: Author's calculations using merged data from education records and earnings data from the LEHD. Sample is 10% of graduates from the University of Texas or public universities in Colorado who have at least three quarters of non-zero earnings and with at least $10,000 of earnings in the calendar year nationally (in-state or out-of-state). National earnings percentile is defined relative to other graduates in the sample with the same residence classification.

Figure 8: Flagship Bias and Differential Attrition by State

A. Bias in Log Earnings

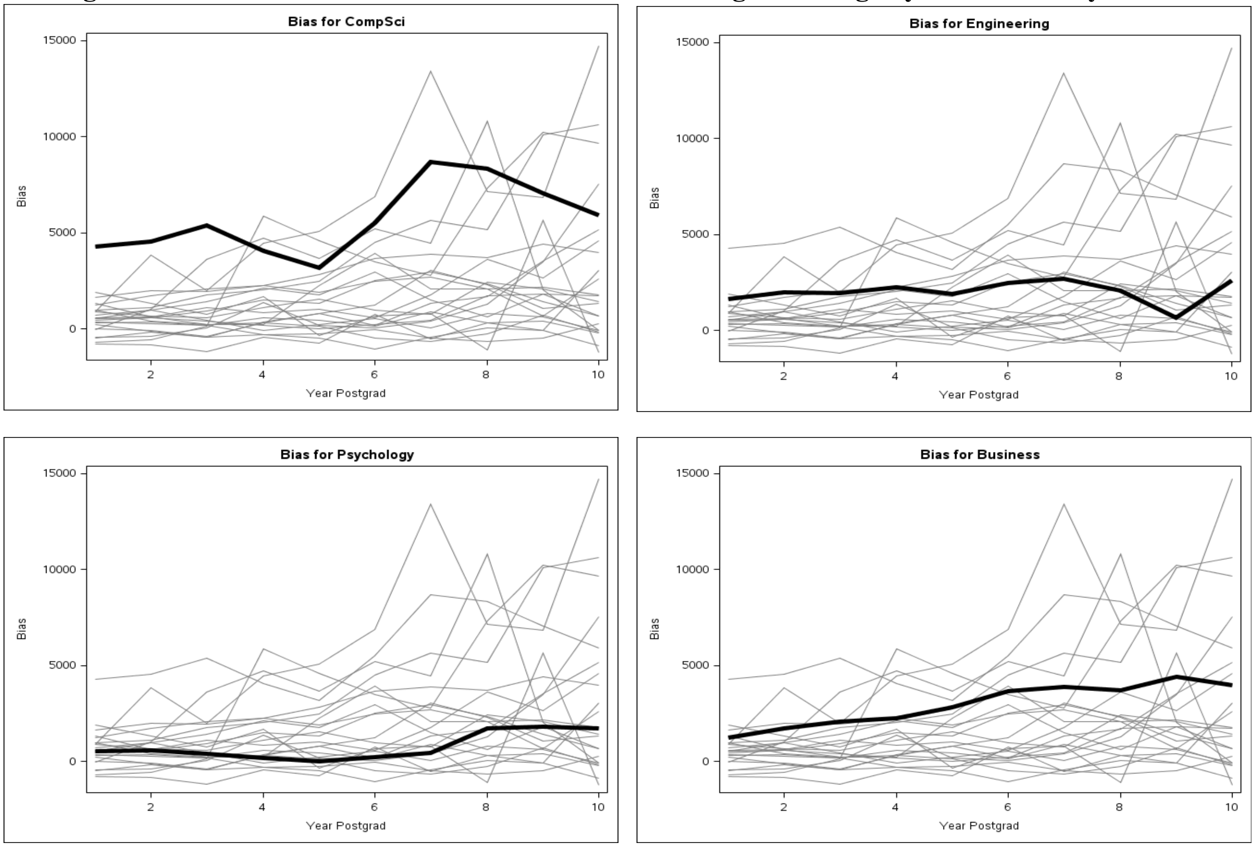


B. Difference in Share with Non-Zero In-State Earnings



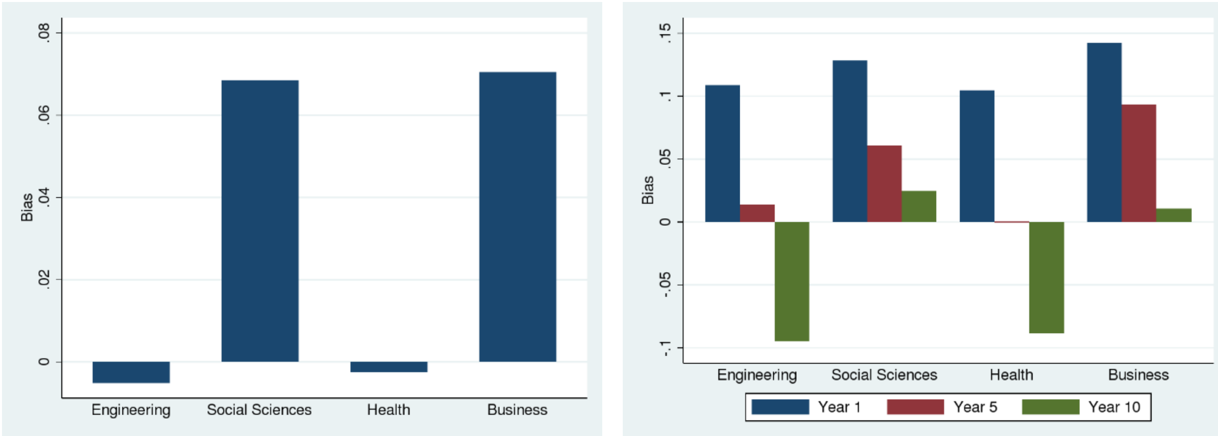
Notes: Bias is estimated by regressing quarterly earnings on a Flagship dummy, fixed effects for each calendar year and quarter, and dummies for male and race (4 categories). This is done for records with non-zero earnings in the state and then nationally. The difference between the flagship coefficients with these two samples is the bias reported in Panel A. These estimates are done separately for cells defined by state and year since graduation. Panel B also reports cell-level regressions, but the outcome is an indicator for whether non-zero in-state earnings is observed. Standard errors clustered by individual.

Figure 9: National minus In-State Difference in Average Earnings by Field of Study



Note: Bias estimates come from comparing average earnings in-state to average earnings nationally.
Sample is 10% of UT graduates who have nonzero earnings in the LEHD.

Figure 10: Bias in Returns to Major



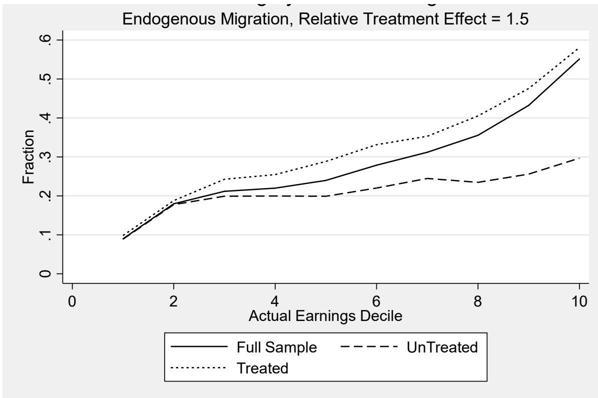
(a) Bias, All Years Combined

(b) Bias, Separate Year Estimation

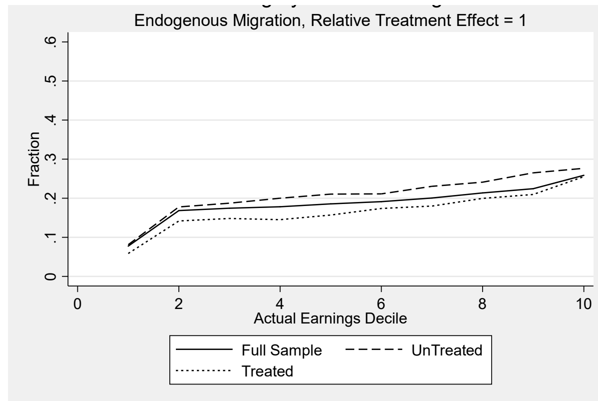
Note: Bias estimates come from the differences in major fixed effects from the regressions described in Section IV.C. For more details, consult the text.

Figure 11. Simulated Share Moving by Earnings Decile

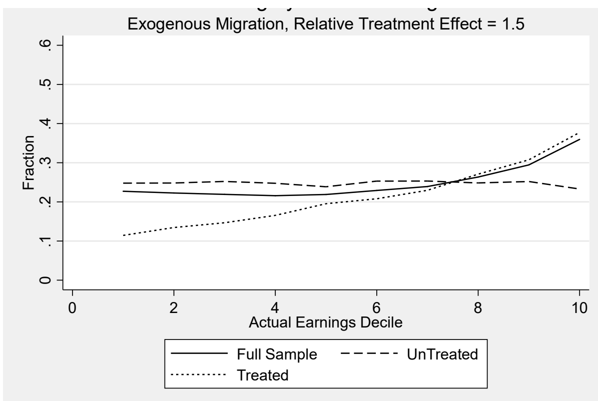
A. Endogenous Migration, $\boldsymbol{\beta_{out} = 1.5\beta_{in}}$



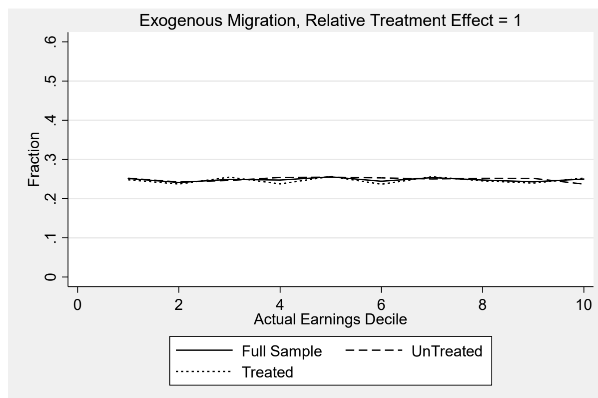
B. Endogenous Migration, $\boldsymbol{\beta_{out} = \beta_{in}}$



C. Exogenous Migration, $\boldsymbol{\beta_{out} = 1.5\beta_{in}}$



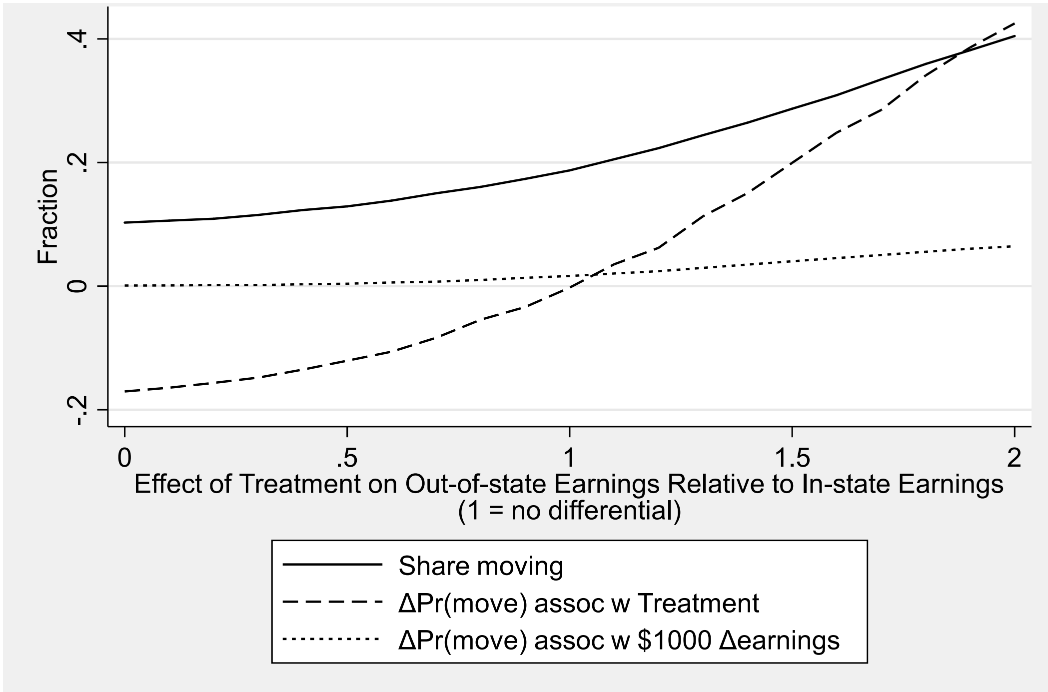
D. Exogenous Migration, $\boldsymbol{\beta_{out} = \beta_{in}}$



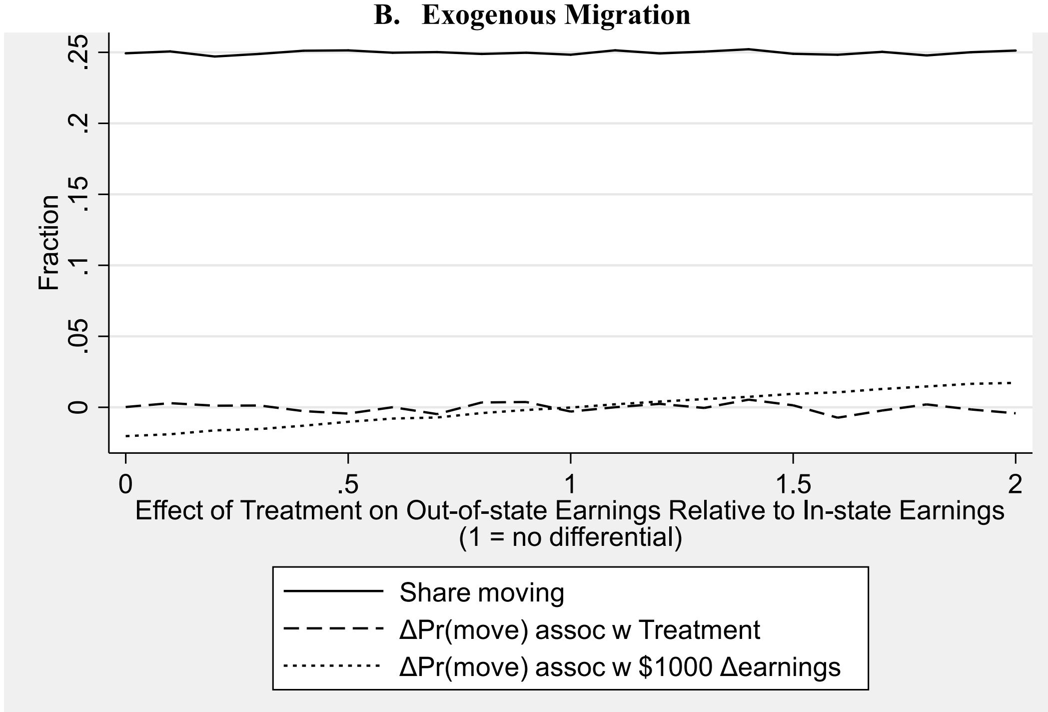
Notes: Results shown from the monte carlo experiments described in Section VI. More detail available in the text.

Figure 12. Simulated Mobility Patterns by Value of Differential Treatment Effect

A. Endogenous Migration



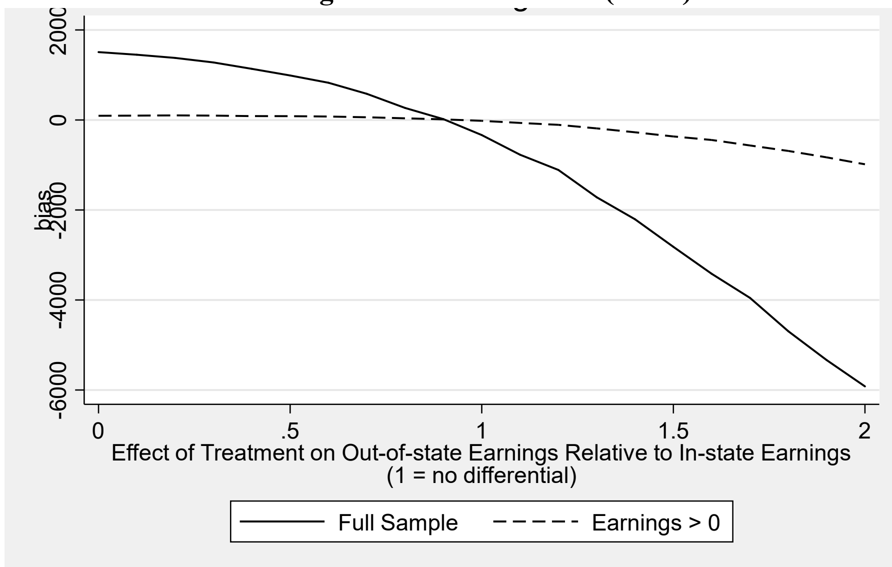
B. Exogenous Migration



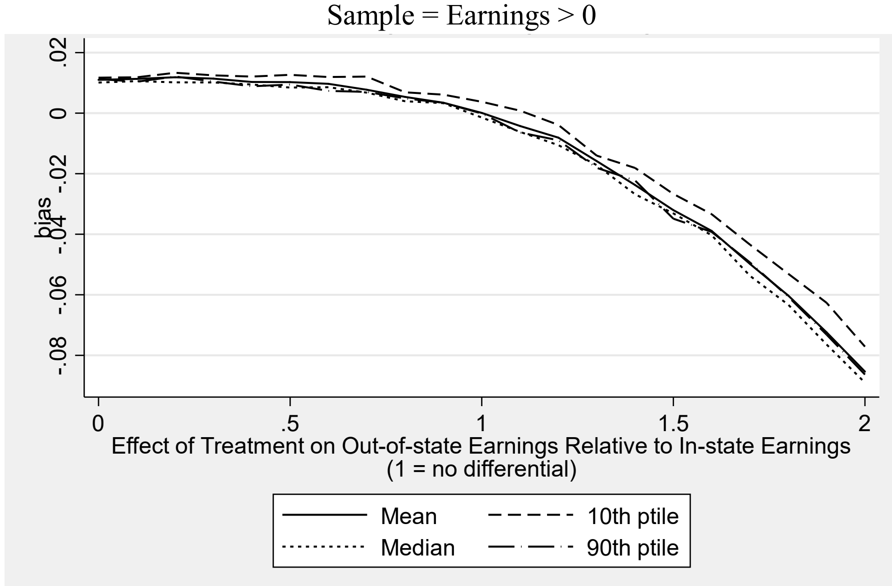
Notes: Results shown from the monte carlo experiments described in Section VII. More detail available in the text.

Figure 13. Simulated Bias in Treatment Effect Estimates with Endogenous Migration
by Value of Differential Treatment Effect

A. Earnings Outcomes in Levels (mean)



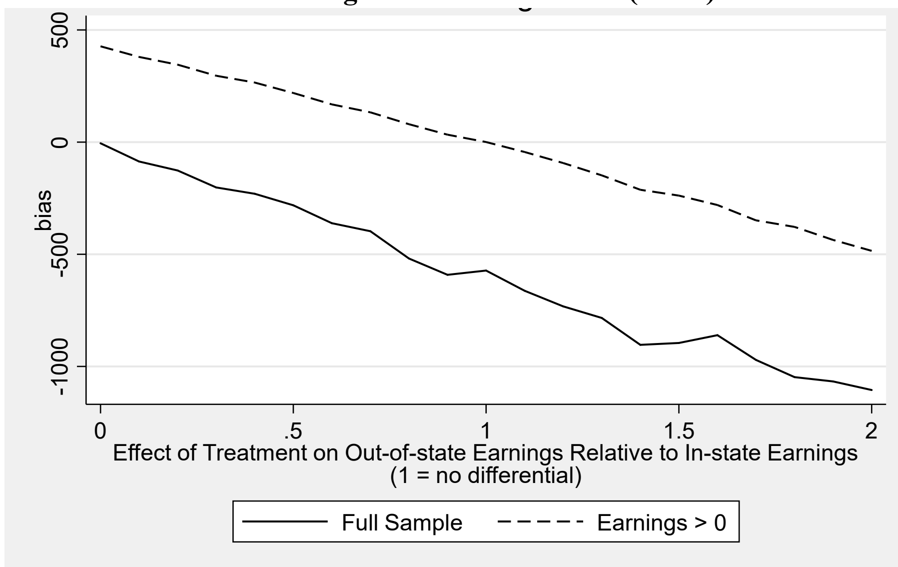
B. Earnings Outcomes in Logs (mean and percentiles)
Sample = Earnings > 0



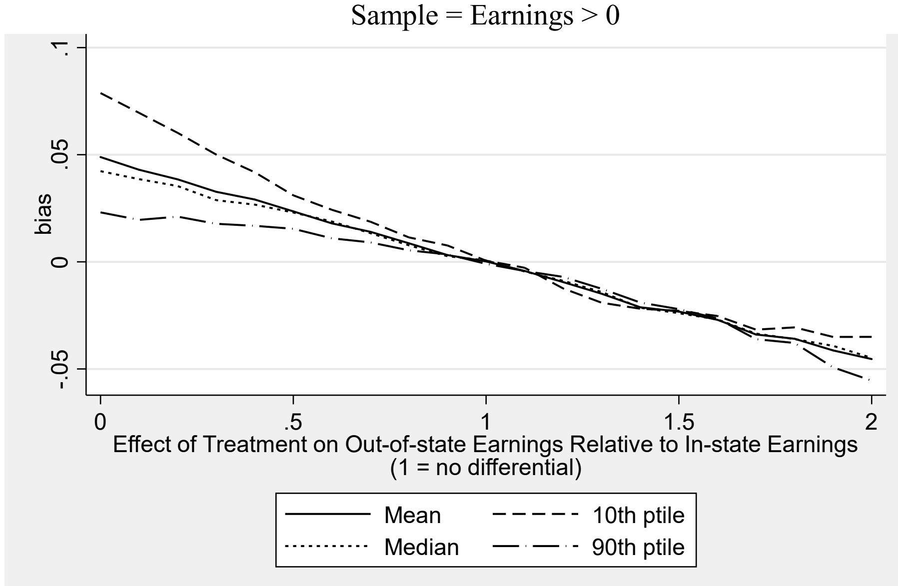
Notes: Bias is calculated as difference between coefficient estimate from OLS regression of observed earnings outcome on indicator for treatment minus estimate from similar regression using actual earnings. Actual uses true earnings outcome and observed uses outcome where earnings is set to zero for individuals that move out of state. Sample restriction of positive earnings is imposed on either actual or observed earnings depending on the specification. Simulations for 100,000 observations use the following parameter values: β_0=$8000, σ_ability=$2000, σ_in=$1000, σ_out=$1000, r_max=$8000, c_max=$3000, β_in=$2000. β_out varies as a multiple of β_in. Migration is endogenous as described in text.

Figure 14. Simulated Bias in Treatment Effect Estimates with Exogenous Migration
by Value of Differential Treatment Effect

A. Earnings Outcomes in Levels (mean)



B. Earnings Outcomes in Logs (mean and percentiles)
Sample = Earnings > 0



Notes: Bias is calculated as difference between coefficient estimate from OLS regression of observed earnings outcome on indicator for treatment minus estimate from similar regression using actual earnings. Actual uses true earnings outcome and observed uses outcome where earnings is set to zero for individuals that move out of state. Sample restriction of positive earnings is imposed on either actual or observed earnings depending on the specification. Simulations for 100,000 observations use the following parameter values: β_0=$8000, σ_ability=$2000, σ_in=$1000, σ_out=$1000, r_max=$8000, c_max=$3000, β_in=$2000. β_out varies as a multiple of β_in. Migration is exogenous (assigned randomly).

Table 1. Recent Articles using Administrative Earnings Records in Postsecondary Education

State	Study	5-year out migration rate
All US states		8.7%
CA	Stevens, Kurleander, Groz (2018)	7.2%
CO	Turner (2015)	13.0%
FL	Hoekstra (2010); Altonji and Zimmerman, (2018); Zimmerman (2014); Bakkes Holzer Valez (2015)	9.1%
KY	Jepson, Troske, Coomes (2012)	7.7%
MI	Bahr, Dynarski, Jacob (2014)	6.1%
MO	Dyke Heinrich Mueser Troske Jeon (2006)	8.4%
NC	Liu Belfield, and Trimble (2015)	8.3%
OH	Minaya and Scott-Clayton (2018); Engbom and Moser (2017); Ost, Pan, Webber (2018)	6.7%
TN	Carruthers Sanford (2018)	8.3%
TX	Andrews, Li, Lovenheim (2016); Andrews and Stange (2016); Cunha and Miller (2014); Denning Marx and Turner (2018)	6.7%

		Emigration rate	
Country	Study	Total	High skilled
Canada	Several	5.5%	6.4%
Italy	Anelli (2018)	5.1%	7.5%
Chile	Hastings, Neilson, Zimmerman (2013)	3.7%	4.0%
Norway	Kirkeboen, Leuven, Mogstad (2016)	4.4%	5.5%
Netherlands	Artmann, Hessel, Oosterbeek, van der Klaauw (2018)	6.1%	8.6%
UK	Belfield, Britton, Buscha, Dearden, Dickson, an der Erve, Sibieta, Vignoles, Walker, Zhu (2018)	8.1%	11.5%
France	Canaan Mouganie 2018		

Sources: Franklin, Rachel S. (2013). "Domestic Migration Across Regions, Divisions, and States: 1995 to 2000" Census 2000 Special Reports. Table 1. August 2003. OECD (2015). *Connecting with Emigrants - A Global Profile of Diasopras 2015.* Table 4.2 total emigration rates and emigration rates of the highly skilled, by country of origin, 2010/2011.

Notes: US out-migration rates pertain to residents of all ages, from 1995 to 2000. Emigration rate is the fraction of citizens at least 15 years old living outside the country in 2010. High skilled refers to those with college degree.

Table 2. Summary Statistics for Full Analysis Sample

	Flagship	Non-Flagship	Difference
National earnings (quarter, include zeros)	13,560	11,080	2480
	(22,320)	(12,050)	
In-state earnings (quarter, include zeros)	8,714	8,867	-153
	(15,890)	(10,980)	
Log national earnings (quarter, zeros dropped)	9.402	9.251	0.151
	(1.005)	(0.892)	
Log in-state earnings (quarter, zeros dropped)	9.368	9.237	0.131
	(0.990)	(0.876)	
Share with no national earnings (true zeros)			
1 year post-grad	0.243	0.182	0.061
10 year post-grad	0.244	0.221	0.023
Share with out-of-state earnings but no in-state earnings (migrants)			
1 year post-grad	0.179	0.097	0.082
10 year post-grad	0.270	0.166	0.104
Observation Count	562,000	1,205,000	1,767,000
Person Count	16,000	36,000	52,000

Notes: Sample includes a 10% random sample of graduates from the University of Texas or public universities in Colorado from 2001 to 2013. This sample of students is merged to quarterly earnings records from the LEHD. Observation counts are rounded to the nearest thousand. All earnings variables are in $2016.

Table 3. Estimates of Effect of Flagship Graduation on Earnings

	No controls				Controls		
	In-state earnings	National earnings	Bias		In-state earnings	National earnings	Bias
	(1)	(2)	(3)		(4)	(5)	(6)
Panel A. Full sample							
Earnings	-719.3	2070	2790		-938.8	1619	2557
	(102.2)	(128.7)			(103.6)	(128.9)	
Panel B. Earnings > 0 (in-state or national)							
Earnings	2746	3465	719		2058	2795	737
	(137.6)	(147.0)			(136.2)	(149.0)	
Log Earnings	0.0867	0.110	0.0236		0.0526	0.0776	0.0250
	(0.0074)	(0.0025)			(0.0074)	(0.0066)	
Log Earn P25	0.0568	0.0737	0.0169		0.0325	0.0491	0.0166
	(0.0022)	(0.0019)			(0.0022)	(0.0020)	
Log Earn P50	0.1043	0.1246	0.0203		0.0689	0.0917	0.0228
	(0.0014)	(0.0012)			(0.0014)	(0.0012)	
Log Earn P75	0.1432	0.1653	0.0220		0.0934	0.1181	0.0247
	(0.0015)	(0.00135)			(0.0015)	(0.0013)	
Log Earn P90	0.1747	0.1983	0.0237		0.1236	0.1543	0.0307
	(0.0020)	(0.0018)			(0.0020)	(0.0018)	

Notes: Dependent variable is quarterly earnings. All models include fixed effects for each calendar year, quarter, and graduation year (but not the interaction) and dummy for being a Texas institution. "Controls" columns additionally include dummies for male and race (4 categories). Standard errors clustered by individual. Panel A includes all observations (n = 1,767,000) while panel B only includes quarter observations for which earnings are non-zero in the state (columns 1,4; 1,090,000 observations) or nationally (columns 2,5; 1,396,000 observations)

Table 4. Estimates of Effect of Flagship Graduation on Earnings, by Year Since Graduation

	Year 1			Year 5			Year 10		
	In-state earnings	National earnings	Bias	In-state earnings	National earnings	Bias	In-state earnings	National earnings	Bias
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Full sample									
Earnings	-1829	-754	1075	-1383	852.7	2236	-204.6	3747	3951
	(106.7)	(114.8)		(160.4)	(180.4)		(248.7)	(310.1)	
Panel B. Earnings > 0 (in-state or national)									
Earn>0	-506.3	-195.5	310.8	1453	1867	414.0	4639	5524	885.5
	(124.5)	(120.7)		(224.0)	(197.9)		(390.5)	(364.4)	
Log Earn	-0.0943	-0.0665	0.0278	0.0389	0.0490	0.0100	0.1447	0.1669	0.0222
	(0.0138)	(0.0124)		(0.0143)	(0.0125)		(0.0158)	(0.0134)	
Log Earn P25	-0.1162	-0.1005	0.0157	0.0124	0.0193	0.0069	0.0929	0.1047	0.0117
	(0.0129)	(0.0115)		(0.0094)	(0.0085)		(0.0084)	(0.0079)	
Log Earn P50	-0.0438	-0.0303	0.0135	0.0565	0.0768	0.0203	0.1435	0.1643	0.0208
	(0.0064)	(0.0059)		(0.0057)	(0.0051)		(0.0065)	(0.0057)	
Log Earn P75	-0.0208	-0.0029	0.0179	0.0752	0.0999	0.0246	0.1884	0.2077	0.0193
	(0.0058)	(0.0054)		(0.0063)	(0.0055)		(0.0081)	(0.0069)	
Log Earn P90	-0.0378	-0.0192	0.0186	0.1140	0.1299	0.0159	0.2711	0.2923	0.0211
	(0.0072)	(0.0067)		(0.0086)	(0.0076)		(0.0111)	(0.0096)	
Observations	57,000	69,000		52,000	68,000		50,000	67,000	

Notes: Dependent variable is quarterly earnings. All models include fixed effects for each calendar year, quarter, and graduation year (but not the interaction), dummy for being a Texas institution, dummies for male and race (4 categories). Standard errors clustered by individual. Panel A includes all observations from graduating classes of 2001-2006 (n = xxxxxx) while panel B only includes quarter observations for which earnings are non-zero in the state (columns 1,4, 7) or nationally

Table 5. Test of Difference in Covariates by Treatment Status and Sample Inclusion

	MALE		WHITE		BLACK		ASIAN		HISPANIC	
	Full sample	Positive In-state Earnings	Full Sample	Positive In-state Earnings	Full Sample	Positive In-state Earnings	Full Sample	Positive In-state Earnings	Full Sample	Positive In-state Earnings
1 Year after	0.0645	0.0683	-0.0212	-0.0330	-0.0256	-0.0243	-0.0033	-0.0017	-0.1740	-0.1895
	(0.0073)	(0.0104)	(0.0052)	(0.0076)	(0.0027)	(0.0040)	(0.0124)	(0.0016)	(0.0051)	(0.0077)
5 years after	0.0645	0.0775	-0.0212	-0.0369	-0.0256	-0.0233	-0.0033	-0.0027	-0.1740	-0.2111
	(0.0073)	(0.0105)	(0.0052)	(0.0076)	(0.0027)	(0.0038)	(0.0124)	(0.0017)	(0.0051)	(0.0079)
10 years after	0.0645	0.0773	-0.0212	-0.0356	-0.0256	-0.0225	-0.0033	-0.0025	-0.1740	-0.2040
	(0.0073)	(0.0106)	(0.0052)	(0.0077)	(0.0027)	(0.0040)	(0.0124)	(0.0017)	(0.0051)	(0.0079)

Dependent variable is covariate listed. All models include fixed effects for each calendar year, quarter, and dummy for being a Texas institution. One observation per person is included. For each outcome, the first column includes all observations from graduating classes of 2001-2013 while the second column only includes bservations for which earnings are non-zero in at least three quarters in the state in that year.

Table 6. Estimates of Effect of Flagship Graduation on Earnings, Lee Bounds

	All Years		1st Year Post Grad		5th Year Post Grad		10th Year Post Grad	
	Coefficier	Lee Bounds	Coefficient	Lee Bounds	Coefficier	Lee Bounds	Coefficien	Lee Bounds
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Earn In-State	3305	[768,6845]	305	[-1605,2697]	2663	[207,5891]	6523	[3551,11080]
	(145.0)		(120.0)		(235.6)		(399.0)	
Log Earn In-State	0.132	[-0.202,0.370]	-0.012	[-0.368,0.212]	0.115	0.195,0.333]	0.232	[-0.066,0.462]
	(0.008)		(0.012)		(0.014)		(0.0159)	
Log Earnings national	0.150		0.007		0.116		0.240	
	(0.007)		(0.012)		(0.012)		(0.013)	
Share trimmed	0.212		0.201		0.215		0.207	

Notes: Sample includes all quarterly observations from graduating classes of 2001-2006 for which earnings are non-zero in the state. Models do not include any covariates or controls.

Table 7. Simulation Results

Sample	Moment	Panel A. Base Simulation $\beta_{out} = 1.5X\beta_{in}$ endogenous migration			Panel B. $\beta_{out} = \beta_{in}$ endogenous migration			Panel C. $\beta_{out} = 1.5X\beta_{in}$ exogenous migration			Panel D. $\beta_{out} = 1.5X\beta_{in}$, $sd\_out\_treat = 1.5Xsd\_out$ endogenous migration			Panel E. $\beta_{out} = \beta_{in}$, $sd\_out\_treat = 1.5Xsd\_out$ endogenous migration		
Sample	Moment	Actual	Observed	Bias	Actual	Observed	Bias	Actual	Observed	Bias	Actual	Observed	Bias	Actual	Observed	Bias
Coefficient on treatment indicator with earnings outcomes																
Full	mean level	2,896	82	-2,814	2,332	1,961	-372	2,727	1,821	-906	3,079	-138	-3,217	2,502	1,530	-973
Earn > 0	mean level	2,315	1,972	-343	1,823	1,803	-20	2,021	1,776	-245	2,510	1,932	-578	1,989	1,806	-184
Earn > 0	mean log	0.248	0.218	-0.030	0.201	0.201	0.000	0.226	0.202	-0.024	0.264	0.214	-0.050	0.216	0.201	-0.015
Earn > 0	p10 log	0.296	0.274	-0.022	0.250	0.254	0.004	0.283	0.260	-0.023	0.300	0.268	-0.032	0.260	0.253	-0.006
Earn > 0	p50 log	0.242	0.209	-0.033	0.195	0.193	-0.002	0.216	0.192	-0.024	0.258	0.205	-0.053	0.209	0.193	-0.016
Earn > 0	p90 log	0.203	0.175	-0.028	0.161	0.162	0.001	0.186	0.162	-0.024	0.232	0.173	-0.059	0.181	0.163	-0.018
Migration and work patterns of sample																
% Move			28.9%			19.0%			24.9%			29.8%			21.1%	
% Don't work			4.4%			4.7%			6.9%			4.4%			4.7%	
Outcome = move																
coefficient on treatment			0.200			0.000			0.002			0.218			0.043	
coefficient on actual earnings (x 1000			0.040			0.017			0.010			0.046			0.025	
Outcome = actual earnings																
coefficient on move			1872			983			569			2271			1441	

Notes: Table reports coefficient estimate from OLS regression of earnings outcome on indicator for treatment. Actual uses true earnings outcome and observed uses outcome where earnings is set to zero for individuals that move out of state. Sample restriction of positive earnings is imposed on either actual or observed earnings depending on the specification. Simulations for 100,000 observations use the following parameter values: β_0=\$8000, $\sigma_{ability}$=\$2000, σ_{in}=\$1000, σ_{out}=\$1000, r_{max}=\$8000, c_{max}=\$3000, β_{in}=\$2000.

Table 8. Monte Carlo Bounding Results

Sample	Moment	Panel A. Point estimates of earnings effects: Actual	Observed	Panel B. Replace the top and bottom X% of positive earnings distribution from control group with zero: Lower	Upper	Panel C. Lee Bounds: exclude the top and bottom X% of positive earnings distribution from control group: Lower	Upper	Panel D. "Tight" Lee Bounds: Compute bounds separately by ventile of ability distribution and re-weight: Lower	Upper	Panel E. "Tight" Lee Bounds: Compute bounds separately by ventile of moving cost distribution and re-weight: Lower	Upper	Panel F. Monotonicity failure: trim X% of control group and Y% of treatment group: Lower	Upper
Coefficient on treatment indicator with earnings outcomes													
Full	mean level	2,896	82	804	1581								
Earn > 0	mean level	2,315	1,972			1,336	2,626	1,874	2,390	1,478	2,512	-1,322	5,249
Earn > 0	mean log	0.248	0.218			0.129	0.290	0.216	0.267	0.153	0.279	-0.143	0.587
			Fraction trimmed from T									45%	
			Fraction trimmed from C			18%		18%		18%		54%	

Notes: Panel A reports coefficient estimate from OLS regression of earnings outcome on indicator for treatment. Actual uses true earnings outcome and observed uses outcome where earnings is set to zero for individuals that move out of state. Sample restriction of positive earnings is imposed on either actual or observed earnings depending on the specification. Panel B replaces bottom or top 17% of observed earnings distribution of control group with zero eearnings to construct lower and upper bound for true estimate, respectively. Panel C excludes the bottom or top 17% of observed earnings distribution of control group from regression. Panel D calculates bound for 20 groups defined by ability and then calculates weighted average of these bounds. Simulations for 100,000 observations use the following parameter values: β_0=\$8000, $\sigma\_ability$=\$2000, σ_in=\$1000, σ_out=\$1000, r_max=\$8000, c_max=\$3000, β_in=\$2000, β_out = 1.5Xβ_in. Migration is

Table A1. Universities Included in Analysis

	Undergraduate enrollment (2009)	spending per fall FTE (2006)	Graduation rate (2006)	Percent admitted (2006)	Composite 25th ptile (2005)	ACT Composite 75th ptile (2005)	Median earnings 10 years	% Instate 1 Year
Texas Institutions								
The University of Texas at Austin	38,168	10,308	77	49	23	29	80,411	80
The University of Texas at Arlington	21,370	5,160	42	69	19	24	67,688	88
The University of Texas at Dallas	9,801	7,738	55	51	24	29	73,592	88
The University of Texas at El Paso	17,205	5,172	29				58,779	84
The University of Texas at San Antonio	25,006	3,969	28	99	18	22	57,938	90
The University of Texas at Tyler	5,051	4,714	40	76	20	25	56,212	84
The University of Texas of the Permian Basir	2,739	4,296	29	86	19	24	53,062	77
The University of Texas Rio Grande Valley	15,947	5,131	33	65	16	20	53,227	91
Colorado Institutions								
University of Colorado Boulder	27,219	8,474	66	88	23	28	67,146	66
Adams State University	2,405	2,786	31	63	17	22	44,170	70
Colorado Mesa University	6,939	3,109	29	82	18	23	50,696	74
Colorado School of Mines	3,672	9,861	68	80	25	29	100,204	55
Colorado State University-Fort Collins	22,221	6,625	64	86	22	26	59,186	75
Colorado State University-Pueblo	4,812	3,511	31	96			48,353	75
Fort Lewis College	3,770	3,966	32	73	18	23	48,818	60
Metropolitan State University of Denver	22,837	3,107	24	83	17	22	54,165	88
University of Colorado Colorado Springs	7,007	4,738	42	62	21	25	57,221	78
University of Colorado Denver	13,246	16,219	36	69	19	25	61,576	83
University of Northern Colorado	10,290	4,352	49	76	19	24	51,355	81
Western State Colorado University	2,237	4,547	31	62	18	23	49,127	59

Source: Authors' analysis from IPEDS data center and https://lehd.ces.census.gov/data/pseo_beta_viz.html. Notes: UT Rio Grande Valley was formed by combining UT Brownsville and UT Pan American in 2013. University of Colorado Denver includes a large medical school, greatly increasing its estimated spending. Median earnings are for graduates who earn at least the annual equivalent of full-time work at the prevailing federal minimum wage and have at least three quarters of non-zero earnings in the reference year.