Plugging the [Brain] Drain

Evidence from online resumes of college graduates

Nicholas Martens
Becker Friedman Institute for Economics
University of Chicago

November 28, 2025

Most recent version available here

Abstract

In the United States, regional brain drain of college graduates is reversible. I link college graduates to sub-state origin and destination labor markets and their colleges using LinkedIn profiles. Roughly half of all college graduates who leave their high school or college local labor market will return within a decade. Positive labor demand shocks in the origin labor market increase the probability that a graduate eventually returns. In both the high school and college labor market, a 1 percent labor demand shock increases the probability of return migration by 1.2 percent. Responsiveness to local labor demand shocks vary based on where someone went to college, and whether they were in-state, transferred, or graduated by age 24. In-state students, old students, and graduates of public institutions are more likely to return.

Acknowledgments
I am indebted to Evan Rose, Ana Reynoso and Brad Hershbein, who encouraged me to aim high with my undergraduate thesis and provided many helpful comments. Additionally, I am grateful to Robert Manduca, Kathleen Bolter, Shane Reed, Nathan Sotherland, Nick Montgomery, Kathryn Dominguez, John Bound, Jimmy Brancho, and conference participants for advising and comments. Gregory Wright and Ian Seyal at the Brookings Institution's Workforce of the Future Initiative generously provided access to the LinkedIn data, while Ana Reynoso shared data storage resources. The Population Association of America supported this work through the NextGenPop Fellowship, which included presenting at the 2024 and 2025 Annual Meetings.

1 Introduction

Over the last 40 years, college graduates increasingly clustered in high-wage, productive metropolitan areas, leading to spatial sorting of the American labor force (Diamond and Gaubert 2022). Spatial sorting raises the specter of a bifurcated nation, with prosperity contained in a few regions and the remainder consigned to distressed labor markets with anemic growth in wages and employment (Moretti 2012). Place-based policies potentially compensate for spatial sorting by redistributing resources to distressed places, though their efficacy remains hotly debated among economists (Glaeser and Gottlieb 2008; Kline and Moretti 2014; Bartik 2020). A place-based policy's ability to improve aggregate welfare hinges on the distribution of workers' idiosyncratic preferences and workers' willingness to migrate as economic conditions change (Gaubert et al. 2025). Place-based policies that stimulate a local labor market might draw additional workers, bringing fiscal benefits that offset project costs. Without causal estimates of the relationship between migration and economic conditions, standard economic impact models might underestimate the benefits of place-based policies. Measuring workers' willingness to migrate poses empirical challenges, with few American datasets tracking individuals' residential decisions over time (Sprung-Keyser, Hendren, and Porter 2022).[1] Resumes on LinkedIn (LI) often include detailed individual characteristics and geocoded resume entries, enabling measurement of longitudinal residential decisions.[2] To mitigate potential selection issues, I focus on workers who report completing a bachelor's degree between 2000 and 2013 at an American college. Using a Bartik instrument to identify quasi-random variation in local labor demand, I estimate the causal effect of local labor demand on migration for college graduates, which represents a potentially important margin for cost-benefit evaluation of place-based policies.

To ensure consistent measurement, I focus on migration within 10 years of college graduation. Return migration is a common feature of college graduates' migration trajectories: roughly half of college graduates will return to either their high school or college labor market, conditional on leaving both after graduating. I divide return migration into three mutually exclusive channels. For workers who completed high school and college in different labor markets, return migration could occur to either labor market. Workers who

1. Until recently, most research on migration relied on repeated cross-sections from the ACS and Decennial Census samples (Diamond and Gaubert 2022; Winters 2020; Wozniak 2010; Bound et al. 2004). In both datasets, most granular geography describing birthplace is state, with no additional variables characterizing the location of high school, college, or other predictors of local ties (Zabek 2024). Sprung-Keyser, Hendren, and Porter 2022 links origins and destinations, but only includes educational attainment for individuals who responded to the American Community Survey (ACS). Some national surveys collect migration data, though LI sample sizes are scales of magnitude larger, allowing for greater geographic coverage. The Panel Study of Income Dynamics restricts access to location data. Likewise, the National Longitudinal Survey used in Groen 2004 includes a migration distance variable, though data isn't as reliable as state and county of residence variables. The Census Bureau's new Post-Secondary Employment Outcomes (PSEO) dataset offers promise, though coverage varies greatly by state and sample sizes remain miniscule (Conzelmann et al. 2022). For example, the only institution in Michigan included in PSEO is the University of Michigan–Ann Arbor.

2. Conzelmann et al. 2023 presents the first application of LI data to migration, relying on summary data by institution, rather than individual-level microdata linking high schools, colleges, and post-graduation outcomes. Section 2 discusses the LI's empirical applications of in greater detail.

completed both in the same labor market can only return to a single labor market. Within each channel, I measure the migration response to two shifts in local economic conditions: an endogenous shift in employment levels and a plausibly exogenous Bartik-style labor demand shock. I construct the shock by subtracting the actual employment level in a local labor market from a prediction that assumes each local industry grew at the same rate as the national leave-one-out industry over the previous year. By using leave-one-out national industry, I avoid introducing mechanical bias into the instrument (Borusyak, Hull, and Jaravel 2022). The instrument predicts state employment growth if each local industry mirrored the shifts in the leave-one-out national industry, removing local supply factors (Bartik 1991). Section 3 describes the empirical strategy in greater detail.

In general, the migration responses comport with Blanchard and Katz 1992: positive labor demand shocks boost in-migration by increasing the probability of return migration, even after controlling for college graduation cohort, institutional characteristics, occupational characteristics, and high school state. For workers who attended college and high school in different labor markets, a 1 percent labor demand shock in the college labor market increases the probability of return migration to that labor market by 0.41 p.p. (1.23 percent). Keeping the same sample but switching to the high school labor market, return migration rises by 0.21 p.p. (1.25 percent) in response to a 1 percent labor demand shock. For workers returning to the high school labor market, the college labor market is a substitute. Relative to a worker whose college experiences no shock, a worker whose college *and* high school labor markets experience positive labor demand shocks halves their probability of returning to the high school. The converse does not hold, suggesting the returners to the college labor market see the high school labor market as a complement rather than substitute. The migration response is more modest for workers who completed high school and college in the same labor market: return migration increases by 0.13 percentage points (0.28 percent) when the high school/college labor market experiences a 1 positive shock to labor demand.

The migration response predicted in Blanchard and Katz 1992 is not evenly distributed among all college graduates. Returners inframarginal to the labor demand shock are disproportionately in-state graduates. For the average in-state graduate of a Regional Public University (RPU), the response to a 1 percent labor shock in the college labor market is approximately 10 percent larger than it is for in-state Public Flagship (PF) alumni. In-state status figures prominently into the migration response to a shock in the high school labor market Groen 2004. Among both RPUs and PFs, in-state graduates are almost twice as likely as out-of-state students to return migrate to the high school labor market, Workers who complete their degree by age 24 are 7.6 percentage points less likely to return to their college labor market compared to someone who graduates at age 25 or higher, assuming the college and high school labor market differ. The same workers are 7.4 percentage points more likely to return to their high school labor market relative to people completing bachelor's degrees later in the life cycle. The magnitude on the traditional student coefficient is similar to the effect of attending an in-state institution on return migration.

I contribute to the literature characterizing the workers who respond to local labor demand shocks, which is integral to assessing the efficacy of place-based policies (Gaubert et al. 2025). The seminal model in Blanchard and Katz 1992 hinges on the population rapidly adjusting to local economic conditions through migration. The original model is agnostic to who moves, treating all workers identically regardless of race, educational attainment, or any other characteristics. Subsequent work demonstrates substantial variation in mobility for different segments of the labor market (Hershbein and Stuart 2024; Yagan 2019; Notowidigdo 2020; Bound and Holzer 2000). Even among the particularly mobile college graduate population, I find heterogeneity based on where they went to college and whether they stayed in-state, transferred, or graduated by 24 (Winters 2020; Groen 2004). The heterogeneity in migration responses based on in-state status and type of college suggest the fiscal implications for economic development might vary by individual institution. The results offer an explanation for the dramatic variation in post-graduation spatial distributions between alumni of different colleges (Conzelmann et al. 2023), which potentially contribute to differences in post-graduation earnings across institutions (Chetty et al. 2020; Mountjoy 2024). By altering the migration patterns of native college graduates, a local labor demand shocks also change the fiscal externalities of higher education funding. When a public college graduate leaves their native state for a more prosperous destination, native state taxpayers lose out (Strathman 1994; Trostel 2010). Despite subsidizing the graduate's education, another state will enjoy the graduate's fiscal benefits, such as higher income tax receipts and diminished consumption of government services (Moretti 2003). I clarify which colleges credential the workers most willing to return to their origin labor markets, counteracting spatial sorting. As state governments contend with spiraling fiscal and demographic pressures, the clarification could help state governments maximize their returns to higher education.

Data limitations constrain longitudinal exploration of migration in the United States. I construct a panel linking people to origins, destinations, and educational institutions with greater geographic granularity than standard public datasets, richer individual detail on educational experiences than federal administrative data, and broader geographic coverage than state administrative data. College graduates respond to improvements in local economic conditions where they completed high school and college by returning at higher rates, suggesting migration represents a potentially important channel for recouping the costs associated with place-based policies.

2 Data

A LinkedIn profile that includes a high school, college, and some post-graduation work experience tells a story missed by standard survey data like the American Community Survey or even administrative data from the IRS. A small but growing literature uses online resume data to advance the economics literature (Kim, Kim, and Park 2025; Curtis, O'Kane, and Park 2024; Dorn et al. 2025; Gortmaker, Jeffers, and Lee 2024; Evsyukova, Rusche, and Mill 2025; Berry, Maloney, and Neumark 2024; Conzelmann et al. 2022). LI links degrees to educational institutions and employment to spatially-defined labor markets.

The two most common methods for exploring migration, ACS microdata and Decennial Census samples, suffer from limitations (Sprung-Keyser, Hendren, and Porter 2022). The ACS asks respondents about the field and degree obtained, but not the institution. The five percent Decennial Census samples remove any possibility of sub-state analysis, since states are the most granular geography describing birthplace and residence in the microdata. Like the ACS, the samples from Decennial Census also lack data on the characteristics of baccalaureate institutions. By contrast, LI enables analysis of sub-state migration *and* characteristics of post-secondary institutions.

2.1 Sample Construction

Revelio Labs scraped public profiles on LI, extracting semi-structured text (Dorn et al. 2025). After applying proprietary algorithms to clean and parse into a structured dataset, the Brookings Institution's Workforce of the Future Research Initiative purchased access to the data.

The raw data contains separate files for positions and educational attainment. I merged the position and education files on unique user IDs to created a panel of residential history, uniquely identified by a user id column in LinkedIn profile data. Revelio provides the raw strings from a user's educational attainment section, so I clean the raw strings and match to high schools and colleges in IPEDS.

I relied on the US Department of Education's National Center for Educational Statistics for characteristics of educational institutions. I obtained data on private and public secondary schools through the 2020-21 Education Demographic Geographic Estimates program. The public school file included enrollment by grade. I excluded any institutions with 0 enrollment in grade 12, namely elementary and middle schools. For post-secondary institutions, I used the 2021 Institutional Characteristics Complete Data File includes geodata, enrollment, completion, public status, and myriad other useful variables at the branch (not system) level. I excluded schools that did not offer a bachelor's degree according to the 2021 data. The complex administrative structure of many public higher education institutions leads to many potential codes for match. I chose unit-level identifiers so that I could analyze heterogeneity within state university systems, which is not possible in some states with OPE IDs (Chetty et al. 2017; Conzelmann et al. 2023). Appendix Section 6.1 contains a more detailed description of the data cleaning procedure.

Not all LI profiles contain the same quantity of data (Dorn et al. 2025). Culling the dataset of profiles with limited information improves the quality of the remaining observations. Most importantly, users must include a high school and a college in the United States. Less than ten percent of college graduates on LI include a high school. According to Table 1, the subset of college graduates who include a high school on their LI profile tend to be younger, graduating from a college in the Northeast, and private institutions relative to the college-educated labor force as a whole. For the current analysis, I impose the following restrictions:

- Baccalaureate institution listed in profile
- High school name listed in profile
- At least one job after college graduation with a location listed
- Graduated from college between 2000 and 2013
- Lived somewhere besides the high school and college labor market after graduating from college

The first three requirements are practical constraints. I cannot measure origin labor markets without some indication of where it might be. Similarly, I need post-graduation residence history to measure any migration. Most broadly, the college graduation year requirements limit (but do not eliminate) concerns about selection among LI users.

The subset of older workers who choose to create a LI profile are not a random sample of the college-educated labor force. Older LI users return migrate at lower rates than the broader college graduate population. Using the PSID, Chan, O'Regan, and You 2020 finds 23 percent of college graduates return to their parents' metropolitan area after graduating college.[3] Given the possibility of family relocating after their child completes high school, I consider 23 percent a suitable upper bound for return migration in the LI sample.[4] Figure 1 shows how return migration to either high school or college labor market varies by college graduation cohort among LI users. About 10 percent of LI users who graduated college in the 1980s returned to their high school or college labor markets. The share gradually increases for more recent college graduation cohorts, converging on the 23 percent estimate in Chan, O'Regan, and You 2020 throughout 2000s and early 2010s cohorts. The return migration probabilities among LI users who graduated after 2000

3. The birth cohorts in the PSID sample employed by (Chan, O'Regan, and You 2020) closely match the current age distribution of the labor force. The oldest respondents were born in 1962 and the youngest in 1999.

4. For example, consider someone who graduates high school in Michigan. While the individual is in college, their family moves to Tennessee. With LI data, the only possible destination for return migration would be Michigan. Yet the same mechanism (social ties accumulated in childhood) might motivate the worker to move to Tennessee. PSID measures the residential locations of parents and children, so the child would only count in the 23 percent figure in Chan, O'Regan, and You 2020 if they relocated to Tennessee after college includes.

approximates the benchmark in Chan, O'Regan, and You 2020. In other words, younger LI users reflect the broader college-educated labor force as LI evolved from a supplement to existing social networks into a rite of passage for recent college graduates.

The lack of return migration among older LI users probably reflects the strategic considerations of individual users (Guillory and Hancock 2012). The older workers who joined LI in the middle of their careers probably saw it as a chance to shore up their professional network. Given the social ties accumulated throughout education, people who work in their high school or college labor market likely saw a LI profile as less useful for networking on average. LI's value as a social networking tool diminishes as workers age, especially if the workers already have deep social ties in their labor market. Rather than trawling the internet for potential connections, older workers rely on professional ties developed through work or educational experiences. Working in the high school or college labor market furnishes workers with a bevy of social ties, only diminishing reliance on LI.

I remove college graduation cohorts after 2013 to limit the threat of measurement bias in later cohorts. Measuring migration patterns shortly after college runs the risk of omitting many subsequent returns. The risk approaches 0 after 20 years according to Figure 6. Figure 6 shows that the overwhelming majority of return migration occurs within 10 years of college graduation, even among older workers. In practice, premature measurement of return migration would likely downwardly bias coefficients, since potential returners might not be recorded as such. While a 20 year cutoff imposes a stringent constraint on measurement bias, the 10 year cutoff substantially increases the sample size due to LI's age distribution. Thus, I subset to college graduation cohorts prior to 2013, the last year in which users could possibly record a decade of post-graduation residential history. For all cohorts, I consistently measure return migration by only reporting returns within the 10 years following college graduation.

2.2 Validation

The data cleaning process includes matching strings of high school and college names to lists of institutions from NCES and IPEDS respectively. No string match perfectly captures everything, and millions of observations make hand-coding implausible. Nonetheless, Figures 2 and 3 suggest the string matching appears to pick up more than noise. Section 6.1 in the Appendix describes the construction of the dataset in greater detail.

Baccalaureate institution names are easily verifiable and (with rare exceptions) unique across the United States. Importantly, baccalaureate institutions are relatively sparse compared to high schools, with secondary schools out-numbering baccalaureate institutions roughly 15 to 1. Figure 2 suggests the string matching procedure for college name offers a reasonable match much of the time. The spatial patterns in the analytical sample mirror the results in Conzelmann et al. 2023, despite their reliance on LI summary data rather than microdata. Conzelmann et al. 2023 provides the first mostly complete picture of spatial patterns among American higher education institutions, dramatically improving on the incompleteness of PSEO. Against the best available benchmark for institution-level migration patterns, the analytical sample holds up well.

Matching on high school names posed a larger challenge. Many high schools across the United States share names.[5] To improve the quality of matches, I match college string names to institutions prior to high schools. In cases of duplicate high school names, I select the high school closest to the college listed on the profile. Most students in the United States attend a college near their home (Alm and Winters 2009; Acton et al. 2024). The matching of high schools and colleges on LI profiles closes a gap in the literature, but the novelty of linked origin-destination data limits the possible avenues for robustness checks. Currently, the only high school-level migration data is restricted-access: administrative data from higher education systems, linked ACS and IRS records, and the Facebook data in Koenen and Johnston 2024. As a result, Figure 3 benchmarks data against an admittedly imperfect measure: the in-state share of incoming freshmen, averaged by cohort size.[6] Figure 3 suggests the high school string matching largely replicates the in-state student composition from IPEDS.

2.3 Characterizing Return Migration

College graduates' sensitivity to labor market conditions is well-established (Hershbein and Stuart 2024; Diamond and Moretti 2021; Cajner, Coglianese, and Montes 2021; Yagan 2019; Diamond 2016; Wozniak 2010; Bound and Holzer 2000). Relative to less-educated workers, college graduates are more likely to leave when their labor market experiences a negative labor demand shock (Notowidigdo 2020; Wozniak 2010; Bound and Holzer 2000). Given the variation in labor market strength across the United States, the clustering of college graduates in stronger labor markets comes as no surprise (Conzelmann et al. 2022; Diamond and Gaubert 2022). Figure 4 measures the net flow of college graduates in local labor markets and states, by institutional grouping. Every labor market sends a certain number of high school graduates to college. Some of these college graduates leave, some will not, and often people will move to labor markets where they did not receive a degree. Regardless, the labor market serves as a destination for a certain number of college graduates. The map simply divides the number of college graduates originating in the labor market by the number residing there after graduation. Purple areas donate college graduates, while

5. For example, the United States contains 89 Trinity Lutheran High Schools and 74 Central High Schools. Nearly one-quarter of American high schools (10,988 out of 45,658) share a name with at least one other school.

6. Since IPEDS only requires reporting in even-numbered years, I calculate the in-state shares based the even years between 2000 and 2020

green areas receive more graduates than originate there. In general, rural areas donate college graduates to urban areas, reflecting graduates' higher utility in urban areas even after accounting for higher living costs (Diamond and Gaubert 2022). Apart from urbanization, college graduates tend to leave the Midwest and Northeast, with Chicago, New York City, and Boston notable exceptions. Even large labor markets like Philadelphia and Detroit donate college graduates to other parts of the country. At the state level, the hollowing of the nation's interior seems more obvious. Outside of heavily urbanized states like Massachusetts, Illinois, and New York, internal migration overwhelmingly benefits the western and southern parts of the United States.

Figure 4 is agnostic about whether people who originate in a labor market end up there. Figure 5 explores the intersection of origin and destination decisions. Each segment of the bar measures a different mutually exclusive migratory relation with the origin labor market: never leaving, leaving and returning, or never returning. Importantly, students in the "Never Left" category can attend college outside their origin labor market or state, but they return immediately after obtaining a bachelor's degree. By construction, the shares of people never leaving their state will be higher than the shares of people never leaving their labor market because states exceed labor markets in size.[7] RPU graduates are more likely to never leave their origin labor market and state, corroborating Hershbein and Stuart 2024. Both PFs and RPUs retain more graduates than Private institutions, likely reflecting the higher concentration of in-state students at public institutions.

If a college graduates returns to their origin geography, they probably do it shortly after college graduation. Figure 6 shows how the cumulative probability of returning to origin labor market varies by years since college graduation. The first five years after graduating college play a crucial role in shaping long-run residential decisions (Wozniak 2010). More than 75 percent of college graduates in the sample return to their origin labor market by the fourth year since graduating college. Next, I subset to people with at least thirty years of post-graduation migration history. The orange line demonstrates that return migration shortly after college graduation is not simply a feature of the over-representation of younger workers on LI. Since Figure 6 plots a cumulative function, the lower levels on the orange line reflect that fact that some return migration still occurs later in life. Nonetheless, even among college graduates with at least 30 years of work experience, more than 75 percent of return migration occurs within the seven years after college graduation. The youthful skew of the LI sample remains useful for studying migration patterns because the decision to move rarely occurs more than a decade after college graduation.

Figure 7 suggests people do not return migrate to origin locations to insure against adverse economic shocks. If people left labor markets with employment rates to return to high school or college labor markets

7. ...with a couple exceptions. For example, Delaware and the Philadelphia labor market or West Virginia and the Washington labor market.

with high unemployment rates, the slopes of the lines on Figure 7 would be negative. In the parlance of Yagan 2014, the return migration is not "directed". While labor demand shocks increase the probability of return migration,[8] college graduates often enjoy a choice set of labor markets with similar economic outlooks. For many migrants, return to the origin labor market is not a response to persistently anemic growth in whatever destination they chose after graduating, but relative improvement in their high school or college labor market. College graduates enjoy a wide "radius of economic opportunity", so their high school and college labor markets compete with dynamic high-wage destinations (Sprung-Keyser, Hendren, and Porter 2022; Diamond and Gaubert 2022).

If someone attended high school and college in different labor markets, return migration operates through two potential channels: either returning to the high school *or* college labor market. The high school and college return probability are mutually exclusive. The sum of the channels approximates the return probability for people who attended both high school and college, implying linear returns to the accumulation of social ties in high school and college. Return migration to the college labor market occurs more frequently than returns to high school. Institutions of higher education stabilize local economies, dulling the effect of negative economic shocks and powering growth (Winters 2020). Return migration to Public Flagship labor markets lags Private institutions and RPUs.[9] Measurement error likely biases the return through the college labor market upward. Users enrolled in graduate school at an institution in the same labor market and working a part-time job might appear to be "return migrating", even though their longer-run residential choices probably include relocation from a college town. Twenty-two percent of workers in the sample return to their high school labor market.[10] Using the PSID over a similar age cohort, Chan, O'Regan, and You 2020 estimates the cumulative probability of return migration to the parental labor market at 23 percent. The mean dependent variable in Tables 3 and 4 are lower, though measurement error potentially bias the coefficient downward. Parents might relocate to a different destination, perhaps retiring to a sunnier locale. As a result, co-residing with parents might not occur in the high school labor market.

Return migration often occurs in the decade after completing a bachelor's degree, hence the reason the main specification in Section 4 uses a binary indicator for return within 10 years as a dependent variable. Return is quite common; 12 percent of college graduates will return to either their high school or college labor market. Conditional on living in multiple labor markets after college, the probability of return migration

8. Section 4 shows that positive labor demand shocks increase the probability of return migration for the channels represented by the blue and green lines on the plot.

9. In many cases, state governments established public flagships to serve the entire state, while RPUs might serve narrower sub-state regions. As a result, state governments often situated PFs in central locations, rendering the college towns isolated in the eyes of many college graduates.

10. According to Table 3, 17 percent of people who completed high school and college in different labor markets will return to the high school labor market ($n = 396,546$). Among those who completed college and high school in the same labor market ($n = 77,695$), 46 percent return. The weighted average is $\frac{(0.17 \times 396,546)+(0.46 \times 77,695)}{396,546+77,695} = 0.22$

rises. For people who completed high school and college in different labor markets, 33 percent return to their college labor markets and another 17 percent return to the high school labor market. Section 4 will explore how return migration through each of the channels responds to labor demand shocks, based on the empirical strategy outlined in Section 3.

3 Empirical Strategy

The next section presents a shift-share instrumental variables approach, accounting for local variation in labor demand with the instrument. Treating local labor demand shocks as exogenous permits a causal interpretation of graduates' response to shocks. The dependent variable is a binary indicator for whether an individual lived in their high school labor market at any point in the ten years after graduating college.

3.1 Measuring Labor Demand

To begin, suppose I am measuring economic growth in local labor market i, with employment distributed across $b \in B$ industries. Employment growth across the entire labor market ΔE_i is the sum of growth rates in local industries E_{ik}, weighted by the share of local employment in industry b:

$$\Delta E_i = \sum_{b=1}^{B} \frac{E_{ib0}}{E_{i0}} \cdot \Delta E_{ib} \tag{1}$$

Equation 1 shows that I can decompose the total employment growth in the labor market ΔE_i, with growth rate in each local industry ΔE_{ib} weighted by its share of local employment $\frac{E_{ib0}}{E_{i0}}$. However, equilibrium outcomes like ΔE_i depend on both labor supply and demand factors. Suppose for example employment declined in the automotive manufacturing industry in Detroit. The decline could reflect lower sales of vehicles, leading employers to shift labor demand inward. However, the decline could also stem from retirements in an aging workforce, with insufficient supplies of younger workers to replace them. In other words, labor supply shifting inward. Both labor supply and labor demand factor into employment growth. Hence, the growth in employment ΔE_{ib} consists of two components:

$$\Delta E_{ib} = \frac{E_{ib1} - E_{ib0}}{E_{ib0}} = g_b^* + u_{ib} \tag{2}$$

where g_b^* represents an industry-wide latent labor demand shock and local labor supply factors constitute u_{ib}. Across every labor market, workers in industry b will face g_b^*, though its effect on total employment depends on local supply factors u_{ib}. Yet local supply factors are endogenous. For example, the age structure of the labor force in a local area reflects the area's relative attractiveness as a migration destination over previous decades (Hershbein and Stuart 2024; Wozniak 2010). Since migrants tend to be young,[11] decades of weak in-migration correlates with a history of anemic local labor demand. College graduates demonstrate particular sensitivity to local labor demand, settling in locations with stronger economies when they graduated college Wozniak 2010. The labor supply of the college graduate stock bears the scars (or fruits) of previous economic shocks. Combining Equations 1 and 2 shows that including ΔE_{ib} in a linear model and estimating with OLS would lead to biased coefficient estimates of labor demand's effect on migration:

$$\Delta E_i = \sum_{b=1}^{B} \frac{E_{ib0}}{E_{i0}} \cdot \Delta E_{ib} = \sum_{b=1}^{B} \frac{E_{ib0}}{E_{i0}} \left(g_b^* + u_{ib}\right) \tag{3}$$

Labor supply factors u_{ib} are endogenous, potentially biasing OLS coefficient estimates when annual employment growth (ΔE_{ib}) is the independent variable. Labor supply decisions of college graduates in a local area (u_{ib} in Equation 2) lack the quasi-experimental variation necessary to identify a causal effect. However, latent labor demand shocks across industries at the national-level (g_b^*) exhibit "plausibly exogenous" variation (Bartik 1991).

To avoid biasing estimates of the relationship between labor demand and migration, I construct a shift-share instrument z_i using leave-one-out national industry-level employment growth $\Delta E_{b,-i}$.[12] Identification via shift-share instrument hinges on the assumption that national labor supply components are orthogonal to labor supply components in local areas. In essence, leave-one-out construction keeps local employment shares from Equation 3 but replaces endogenous ΔE_{ib} with national shifts by industry $\Delta E_{b,-i}$:

$$z_i = \sum_{b=1}^{B} \frac{E_{ib0}}{E_{i0}} \cdot \Delta E_{b,-i} \tag{4}$$

$\Delta E_{b,-i}$ proxies for g_b^* in Equation 2, which I cannot observe (Bartik 1991; Borusyak, Hull, and Jaravel 2025). To avoid the shifts affecting the shares, the local industry shares ($\sum_{b=1}^{B} \frac{E_{ib0}}{E_{i0}}$) are fixed at 2000 levels. Time-invariant employment shares, fixed to levels at the beginning of the analytical period, limit one potential channel for generating heterogeneity in the shocks across unobservable characteristics (Borusyak, Hull, and Jaravel 2025). To limit contamination by local labor supply factors, $\Delta E_{b,-i}$ uses leave-one-out construction. No two labor markets will experience the same industry-level shift $\Delta E_{b,-i}$ because I subtract the contribution of labor market i when estimating instrument z_i. So $\Delta E_{b,-i}$ measures the shift in employment for industry b

11. Section 2.3 suggests more than three-quarters of college graduates who eventually migrate do it within ten years of college graduation.

12. Within the context of domestic migration in the United States, Bartik-style shift-share instrumental variables remain a popular identification strategy. In most cases, the dependent variable is the population level (Zabek 2024; Notowidigdo 2020; Amior and Manning 2018; Cadena and Kovak 2016). Only Wozniak 2010 regresses a shift-share labor demand shock on a binary migration-related outcome.

across every labor market *except* i (Autor and Duggan 2003). Leave-one-out construction limits mechanical correlation between the local labor supply shocks and the error term that arises from finite sample bias (Borusyak, Hull, and Jaravel 2022). Thus, estimation of labor demand instrument z_i circumvents labor supply factors u_{ib} by construction (Bartik 1991).

The demand instrument constitutes a prediction of employment levels for labor market i, assuming local employment growth ΔE_{ib} equals national employment growth ΔE_b for all industries $b \in B$. In reality, $\Delta E_{ib} \neq \Delta E_b$, leading to a discrepancy between the actual level of employment E_{i1} and z_i. The difference between E_{i1} and z_i is the labor demand shock. The second-stage regression described in Equation 6 includes labor demand shocks for the local labor markets where an individual attended high school and college. To generate regression coefficients that double as semi-elasticities, I approximate the shocks using logs:

$$\Delta\hat{\theta}_i = \frac{E_{i1} - z_i}{E_{i1}} \approx \log E_{i1} - \log z_i \tag{5}$$

Jensen's Inequality suggests taking the log of z_i might bias the estimates since labor markets with employment dispersed across many industries will have higher $\log z_i$, on average, compared to labor markets with concentrated employment (Borusyak, Hull, and Jaravel 2025). However, semi-elasticity coefficients simplifies comparison to other estimates of migration responses.

3.1.1 Identifying Assumptions

Shift-share instrumental variables offer two paths for leveraging exogeneity: exogenous shares (Goldsmith-Pinkham, Sorkin, and Swift 2020) or exogenous shifts (Borusyak, Hull, and Jaravel 2022; 2025). Capital (particularly in heavy industry) is notoriously hard to move, which subjects the industrial composition of local labor markets to hysteresis. In the long run, hysteresis raises the specter of systematic variation in unobservable characteristics across labor markets with higher shares of certain industries. For example, labor markets with high shares of workers in manufacturing might exhibit distinct age structures or educational attainment patterns from labor markets dependent on other industries. Regardless of the potential confounders, local employment shares by industry are endogenous. Despite endogenous exposure shares, shift-share instruments still offer valid identification (Borusyak, Hull, and Jaravel 2022). Endogenous exposure shares ($\frac{E_{ib0}}{E_{i0}}$) must be paired with exogenous shifts in employment for industries nationwide ($\Delta E_{b,-i}$). The identification strategy hinges on a simple assumption: a share-weighted average of random shifts is as-good-as-random (Borusyak, Hull, and Jaravel 2025).

Valid identification through exogenous requires satisfying the exclusion restriction. Exogenous shifts in national industries only affect probability of return migration through labor demand, with no effects on labor supply. The shifts need not be random with respect to industry (Borusyak, Hull, and Jaravel 2025) Since shares are endogenous, we expect certain industries to concentrate. Local labor markets with high labor supply shocks (measured by the error term) might have different industries than those with low labor supply shocks. However, the national employment shocks should not systematically over-represent the subset of industries in either type of labor market (Borusyak, Hull, and Jaravel 2025). Exogenous shifts in employment levels should not systematically vary between labor market types. For example, shifts should occur at random across high- and low- skill industries. Thus, places specializing in particular industries should have typical values of the instrument. Violations of the condition mean the instrument varies systematically across with respect to supply shocks, defying the exclusion restriction.

Borusyak, Hull, and Jaravel 2025 argue the canonical construction of the shift-share instrument in Bartik 1991 suffers from endogeneity without the assumption of share exogeneity. The instrument in Bartik 1991 is not leave-one-out, meaning the labor demand shift ΔE_b for industry b is identical across all labor markets. Leave-one-out shift-share construction potentially mitigates the bias that Borusyak, Hull, and Jaravel 2025 describe in the Bartik 1991 instrument. Use of leave-one-out in shift-share instruments is akin to use of jackknife IV estimation to avoid the bias of weak instruments with 2SLS (Angrist, Imbens, and Krueger 1999; Borusyak, Hull, and Jaravel 2025). However, Borusyak, Hull, and Jaravel 2025 warns of correlation among the error terms of observations similar to i, which would render omission of i alone insufficient for preserving the exclusion restriction.

3.1.2 Data

Since LinkedIn data only includes metropolitan statistical areas (MSA), a subset of the full universe of core-based statistical areas, I use a blend of state and MSA-level data for estimates. Counties outside of metropolitan areas were grouped together by state, forming pseudo-MSAs which aren't necessarily contiguous. However, all Census-defined MSAs exhibit contiguity. For the rest of the paper, I use the labor market to denote the county-delimited sub-state geography. Figure 11 in the Appendix maps all of the labor markets in the United States.

I used 3-digit NAICS codes for the industries, measuring dozens of industries across labor markets. County Business Patterns data does not cover employment in agricultural production, private household services, or the public sector (Cadena and Kovak 2016). Since the US Census Bureau suppresses County Business Pattern data at small sample sizes, the public data omits many industries with relatively few employees in a particular geography. In a small labor market such suppression can skew the employment shifts using the Bartik instrument. I circumvented the suppressions by using Unsuppressed County Business

Patterns estimated by Eckert et al. 2021.[13] The Unsuppressed County Business Patterns cover 1975-2017, though I only use 1999-2013 data because I restrict to the 2000-13 college graduation cohorts.

3.2 Regression Model

I measure all post-graduation migration behavior, constrained by the time horizon of County Business Patterns data from Eckert et al. 2021. Ultimately, the early extent of the time horizon matters little because half of the LI user sample graduated from college in the last 10 years and the first decade after college graduation heavily influences long-term settlement decisions (Wozniak 2010). Less likely to be constrained by complex familial considerations, Americans exhibit the most mobility in their twenties, when most people obtain post-secondary credentials (Wozniak 2010).

I index observations by the labor market of the high school, denoted by j. Importantly, only the high school needs to be in j; the college can be anywhere in the United States. However, people who stay in the same labor market for high school and college ($j = k$) face a different decision-making process than people who leave their high school labor market for their bachelor's degree ($j \neq k$). People who attend college elsewhere compare the conditions of their home labor market j relative to their college labor market k. Those who stay cannot make such a comparison. Their decision to remain depends on the conditions of the state where they attended both college and high school. Hence I report results for those whom $j \neq k$ separately from those for whom $j = k$. Tables 3 and 2 report results for $j \neq k$, while Table 4 for $j = k$. Equation 6 describes the return rate as a function of local economic conditions in the high school labor market j and college labor market k if $j \neq k$:

$$stay_{ij} = \beta_1 hsLD_{jt} + \beta_2 colLD_{kt} + \beta X'_{jkt} + \phi_M + \alpha_t + \mu_{ijt} \tag{6}$$

where $stay_{ij}$ is a binary indicator for individual i ever residing in their high school labor market j at any time $t \in T$ after college graduation such that

$$stay_{ij} = \begin{cases} 1 & \text{Individual } i \text{ attended high school in labor market } j \text{ and resided in labor market } j \text{ at a time } t \in T \\ 0 & \text{Individual } i \text{ attended high school in labor market } j \text{ and never resided in } j \text{ at a time } t \in T \end{cases}$$

$hsLD_{jt}$ is the labor demand shock at time t in the labor market j containing the high school,

$colLD_{kt}$ is the labor demand shock at time t in labor market k containing the college,

13. Cadena and Kovak 2016 also uses County Business Patterns data to measure labor demand shocks. However, they restrict their analysis to larger labor markets, limiting the need for unsuppressed data.

X_i' is a vector of individual controls: institutional grouping of college, in-state status, transfer status, whether 25+ at college graduation, and major,

ϕ_M measures fixed effects in state M containing high school state j such that $j \subseteq M$,

α_t reports proportional shocks across all states in a given time period, and

μ_{ijt} captures unobservable error.

4 Results

4.1 First-Stage Regression

Labor demand shocks instrument for employment growth (Bartik 1991). In the first stage, I test the strength of the instrument by regressing the full model in Equation 6 on the endogenous variable, annual employment growth. To ensure my coefficients represent semi-elasticities, I approximate employment growth with the year-over-year difference in log employment. Since the model includes two instruments (college and high school labor demand shocks), the first-stage consists of four equations that control for the fit of each instrument:

1. Equation 6 with employment growth in high school labor market as dependent variable
2. Equation 6 with employment growth in college labor market as dependent variable
3. (1) controlling for the fitted values of (2)
4. (2) controlling for the fitted values of (1)

A causal interpretation of the second-stage requires independent variation in both instruments. The instrument should vary with respect to high school labor demand growth, conditional on the component of the college labor demand growth predicted by the first-stage regression. The third and fourth equations described above measure the independent variation in high school and college employment growth respectively. Both first-stage equations were quite similar, with coefficients equal to 0.98 and standard errors equal to 0.003. The partial f-statistic associated with each equation provides a heuristic for assessing the strength of the instrument. Both partial f-statistics exceed 1,000,000, with high school narrowly exceeding college. Similarly, the R^2 for both first-stages was 0.98, suggesting the instrument strongly predicts the endogenous variable.

4.2 Second-Stage Regression

Tables 2, 3, and 4 estimate Equation 6 on two sets of independent variables: endogenously-determined employment growth and plausibly exogenous labor demand shocks. The universe is college graduates who lived somewhere besides the origin labor market for at least one year after college graduation. Return migration to the origin labor market occurs through three distinct and mutually exclusive channels: workers who completed high school and college in different labor markets could return to *either* labor market, or workers completed both high school and college in the same labor market and returned. Tables 2 and 3 provide results for the former, while Table 4 the latter. Thus, the dependent variable in Table 2 is an indicator for whether a worker returns to the college labor market at some point within the 10 years after graduating college. Table 3 measures return to the high school labor market instead, though otherwise mirrors the dependent variable in Table 2. To avoid autocorrelation, Table 4 only presents estimates of Equation 6 with a single labor demand variable from the high school and college labor market.

The coefficients are semi-elasticities, measuring the *percentage point* change in probability of return migration to the origin labor market in response to a labor demand shock. Tables 2, 3, and 4 report estimates using two measures of change in local labor market conditions. For the columns labeled "OLS", the change is annual employment growth in the year prior to college graduation. Supply and demand factors endogenously determine annual growth (Bartik 1991). In the "IV" columns, the labor demand shock in the year prior to college graduation measures change. The labor demand shock is the change in the high school labor market's total employment relative to a prediction based on industrial mix of the labor market. Section 3.1 describes the measurement of labor demand in greater detail.

Column (1) across Tables 2, 3, and 4 simply regresses changes in local labor markets on return migration. Congruent with the model in Blanchard and Katz 1992, the College LD Shock coefficient in Table 2 and the HS LD Shock coefficient in Table 3 are positive, indicating that positive labor demand shocks induce additional migration. Column (2) controls for college graduation year fixed effects, ensuring college graduates are compared with peers who faced identical national labor and local labor market conditions. Column (3) controls for the institutional grouping of the colleges, in-state status, major, post-graduation occupational group, and the prestige of the post-graduation occupation. Groen 2004 suggests in-state status modulates migration status of college graduates, while Conzelmann et al. 2023 demonstrates how institutional characteristics shape migration choices. Column (4) adds more individual controls: transfer status and whether the worker completed their BA by age 24. Transfer students exhibit particularly strong preferences for the locations where they complete bachelor's degrees, reflecting preferences strong enough to induce a transfer (Hillier et al. 2023). Finally, Column (5) includes high school state fixed effects, exploiting variation in labor demand shocks within a state. Given the importance of state governments in organizing and financing higher education, institutional details affecting financial aid availability, institutional quality, and programmatic offerings can vary dramatically by state. High school state fixed effects narrow to a more relevant comparison group by fixing the relative attractiveness of in-state tuition.[14] Students in Rhode Island enjoy in-state tuition at two baccalaureate institutions, while Texans and Californians can choose from dozens of branches across multiple distinct university systems. In turn, colleges shape a graduates' choice set of post-graduation migration destinations (Manduca 2024; Conzelmann et al. 2023). By narrowing to a subset of people with similar constraints on the college choice problem, high school state fixed effects compare college graduates with more similar post-graduation migration choice sets.

Table 2 shows that return migration to the college labor market is particularly sensitive to labor demand shocks. A 1 percent shock to labor demand in the college labor market boosts return migration by 0.41 p.p. (1.25 percent). A 1 percent increase in the labor demand shock is approximately a 1 standard deviation increase. Improvements in the high school labor market might modestly reduce return migration to the college labor market, but the coefficients are indistinguishable from zero.

Labor demand shocks' effect on return migration to the high school labor market is smaller but still positive according to Table 3. An identical 1 percent shock in the high school labor market boosts the probability of return migration by 0.21 p.p. (1.4 percent). Return migration to the high school and college labor markets differ in substantive ways. Improvements in the college labor market reduce return migration to the high school labor market according to Table 3, while the converse does not hold in Table 2. A positive labor demand shock in both the high school and college labor markets reduces return migration to the high school by half relative to a world where the high school labor market experiences a positive shock and college experiences none. The reduction in return migration amounts to approximately 0.1 p.p. under the preferred specification.

The significance and opposite signs on migration semi-elasticities suggest the choice sets differ for inframarginal returners to each labor market. College graduates weighing a return to the college labor market generally choose from a set of locations that excludes their high school labor market. Some might consider the college labor market, which could explain the negative sign on the HS LD Shock coefficient in Table 3, but the lack of significance suggests the tradeoff between high school and college labor market is not salient for the average college graduate returning to the college labor market. By contrast, potential returners to the high school labor market generally don't include the college labor market in their choice set. Holding all else constant, a 1 percent positive shock in the college labor market reduces return migration to the high school labor market by 0.11 percentage points (0.66 percent) As local labor markets improve, college graduates'

14. With notable exceptions for municipal colleges that offer bachelor's degrees like the City University of New York

probability of ever returning to the labor markets where they received a high school diploma and bachelor's degree increase.

Unlike Tables 2 and 3, Table 4 relies on the subset of college graduates who attend high school and college in the same labor market. The point estimates in Table 4 are not as precise: smaller sample sizes lead to larger standard errors. After controlling for year of college graduation, institutional characteristics, occupational characteristics, and high school state, return migration increases by 0.13 percentage points (0.28 percent) when the high school/college labor market experiences a 1 positive shock to labor demand. The migration response is muted compared to the migration semi-elasticities for college graduates who attend college and high school in different labor markets. Note that the unconditional probability of return to the labor market containing college and high school is 46 percent,[15] much higher than the unconditional mean probabilities of return to labor market containing high school only (17 percent) or college only (33 percent).[16] The low sensitivity and high mean dependent variable reflect the distinct preferences and demographics of people who stay in the same labor market for college. Under-represented minorities and socioeconomically disadvantaged students tend to exhibit greater sensitivity to distance when selecting colleges (Acton et al. 2024). Even if they do leave after college graduation, they return at higher rates than their peers return to either high school or college labor markets. The outsize weight placed on proximity to friends and family suggests the universe of college graduates in Table 4 will return regardless of economic conditions, which pushes the semi-elasticity downward.

4.2.1 Bias

Consistent with Bound and Holzer 2000, OLS coefficients tend to be smaller than IV estimates; bias arising labor supply shifts pushes point estimates downward. The discrepancy between OLS and IV results is especially durable for the College LD Shock in Table 2, growing to approximately 0.07 percentage points with the full set of controls in Model (5). By contrast, the gap between OLS and IV for the HS LD Shock in Tables 3 and 4 is 0.02 and 0.03 percentage points respectively. Demographic structure potentially explains the gap in biases between the college and high school labor markets. An aging labor force could bias OLS coefficients upward: net declines in labor supply stemming from retirements and mortality also reflect decades of low in-migration over previous decades. In effect, demographic hysteresis in the labor market will counter the downward bias in labor supply described in Bound and Holzer 2000. Yet on average, labor markets with substantial student populations are younger. Hence, the gap between OLS and IV estimates is larger for the college labor market without demographic factors attenuating the downward bias of labor supply

15. See the penultimate row of Table 4
16. See the penultimate row of Tables 2 and 3

Selection bias could contaminate the coefficient estimates presented in Tables 2, 3, and 4 because LI usage is voluntary and self-reported (Guillory and Hancock 2012; Berry, Maloney, and Neumark 2024). The bias likely pushes coefficient estimates of return migration downward. Older LI users (i.e. cohort graduating college prior to 2000) exhibit idiosyncratic migration patterns, returning to origin labor markets at lower rates than Chan, O'Regan, and You 2020 estimates using nationally-representative surveys. Return migration rates by cohort in Figure 1 increased dramatically among later cohorts, perhaps reflecting LI's spread from white-collar in large cities to college graduates across the entire country. Among younger cohorts, LI users comprise a growing share of the college-educated population, easing the selection bias problem.[17] Importantly, return migration matches other estimates after 2000 (Chan, O'Regan, and You 2020). The cross-sectional composition of the LI user population changed between the 2000 and 2013, increasing the geographic scope and representation across post-secondary institutions.[18]. The introduction of cohort fixed effects in Model (2) reduces the downward bias on the OLS coefficients in Tables 3 and 4.

The regression model in Equation 6 assumes that social ties established in the high school and college labor markets are immobile. In reality, friends and family relocate, so social ties could motivate moves to locations besides the high school and college labor market. The assumption embedded in Equation 6 raises the specter of measurement error: in some cases, a worker's social ties may not be concentrated in the high school and college labor markets. In other words, the regression will not capture the important tradeoff for the worker because their social ties will be concentrated in other labor markets. Measurement error in the spatial distribution of a worker's social ties will downwardly bias coefficient estimates in Tables 2, 3, and 4.

4.2.2 Benchmarking

Wozniak 2010 also uses shift-share to construct labor demand shocks, analyzing how people in their 20s respond to shocks in their state of birth. Wozniak 2010 measures two types of labor demand shocks, aggregate shocks to a state's entire labor force and education-specific shocks. The instrumental variable used to estimate the coefficients in Tables 2, 3, and 4 resembles the aggregate shock.[19] The coefficient estimates that imply that a 1 percent labor demand shock increases the probability of a college graduate residing in their state of birth by 0.29 p.p.[20] Importantly, I measure geography at a more granular spatial

17. Figures 15 and 16 measure the age distribution of the LI sample relative to the college-educated labor force in 2023.

18. For additional details about cross-sectional changes in the sample, compare the second and fourth columns of Table 1

19. The column labeled "Bartik" in Panel A of Table 3 of Wozniak 2010.

20. The main effect measures the migration response of all workers to a 1 standard deviation increase in the size of labor demand shock, while an interaction term measures the response for workers with a bachelor's degree or more. Adding the main effect (-0.016) and interaction term (0.013) yields -0.03. To convert to a semi-elasticity, I divide by the standard deviation on the "Bartik" variable reported in Panel A of Table 2 (0.019). Since Wozniak 2010 uses residence *outside* state of birth as a dependent variable, I flip the signs on coefficients to obtain $0.16 = \frac{0.03}{0.19}$. A 1 percent shock to labor demand increases the probability of residing *inside* the state of birth by 0.16 percentage points. Forty-five percent of college graduates in Wozniak 2010 reside within their state of birth, so a 1 percent labor demand shock leads to a $0.36 = \frac{0.16}{0.45}$ percent increase in the population of college graduates.

level (sub-state local labor market) with a younger cohort of college graduates[21] than Wozniak 2010The estimate from Wozniak 2010 falls between the migration semi-elasticities with respect to high school labor market (0.21) and college (0.41) for people who attended high school and college in different labor markets (Tables 2 and 3).

4.2.3 Heterogeneity

The propensity for return migration varies across sub-groups of college graduates. Note that only semi-elasticity coefficients are in the "HS LD Shock" and "College LD Shock" rows of Tables 2, 3, and 4; the subsequent rows measure level shifts. For non-semi-elasticity rows, the coefficient measures the percentage point change in return migration probability to the origin location (however it may be defined in the specific table) relative to a reference group.

Transfer and non-traditional[22] students are more attached to the college labor market than the high school labor market. Holding all else constant, a college graduate who attended multiple colleges is 0.022 percentage points (6.7 percent) more likely to return to the labor market of their final college when that labor market experiences a demand shock than a college graduate who did not transfer. Table 2 shows that the transfer effect is substantial and precisely estimated. The magnitude on the Transfer coefficient is approximately 40 percent of the increase in return migration to the college labor market associated with in-state status. Workers who complete their degree by age 24 are 7.6 percentage points less likely to return to their college labor market compared to someone who graduates at age 25 or older, assuming the college and high school labor market differ. According to Table 3, the same workers are 7.4 percentage points more likely to return to their high school labor market relative to people completing bachelor's degrees later in the life cycle. The magnitude on the traditional student coefficient is similar to the effect of attending an in-state institution on return migration. The similarity in magnitudes and difference in coefficients suggests workers who graduate by age 24 face a distinct decision problem when selecting a post-graduation migration destination. Traditional graduates typically face fewer familial or financial constraints when selecting their college; they are more likely to be white, male, and socioeconomically advantaged (Bárány, Buchinsky, and Corblet 2023) The comparative lack of constraints prior to college translates into a wider "radius of opportunity" after college (Sprung-Keyser, Hendren, and Porter 2022). By contrast, non-traditional graduates often face steeper distance constraints when selecting a baccalaureate institution. Graduates over 24 are more likely to balance pre-existing familial and employment obligations (Bárány, Buchinsky, and Corblet 2023). The constraints

21. Wozniak 2010 uses the 1950-79 birth cohorts while I use people who graduated college from 2000 to 2013. Note that I don't impose restrictions on the birth, but the 90th percentile with respect to cohort in the LI sample is 1975 and the 10th percentile is 1991 according to the "Primary Sample" column of Table 1.

22. Workers who graduate from college at age 25 or older

lead many non-traditional students to complete their degree at a proximal institution and remain nearby after graduation. Groen 2004 documents the importance of in-state status in shaping migration responses, though other characteristics of a post-secondary trajectory like age at graduation and transfer status affect return behavior in substantive ways.

Figure 17 estimates complier means of various covariates using the Kappa theorem in Abadie 2003.[23] Compliers are college graduates induced to return to the high school labor market by the labor demand shock, so Figure 17 measures selection among in-migrants with local ties (blue circle). Figure 17 measures selection relative to two populations: Leavers and the union of Leavers and Stayers. Since all compilers must live outside the high school labor market before returning, compliers are a subset of Leavers, the college graduates who leave their origin labor market after graduation (represented with an orange triangle). Stayers are college graduates who never leave their high school labor market. The purple triangle represents the full universe of college graduates, with no conditions imposed on post-graduation migration behavior. Compliers are 12 p.p. more likely to attend college in-state, 5 p.p. more likely to complete their bachelor's degree before by age 24, and 2 p.p. less likely to transfer between colleges. With respect to institutional grouping, compliers attend the same mix of colleges as other populations. Likewise, Figure 17 suggests minimal selection on the basis of occupational prestige, suggesting compliers aren't merely workers who "struck out" in other labor markets.[24] By construction, compliers must return to the high school labor market within ten years. Hence, the share of compliers returning within 5 years or 5-10 years mechanically exceeds the share for Leavers, who need not return, and Stayers, who never leave. The college graduates induced to return to their high school labor market by labor demand shocks are more likely to attend in-state institutions and complete their degree by 24, but no different in their occupational or institutional composition.

Figure 8 plots the implied migration semi-elasticities in column (5) by in-state status and institutional group. The differences between RPUs and PFs stands out in Figure 8. The difference might reflect endogenous amenities: RPUs are more likely to be located in rural or isolated regions, while PFs tend to be centrally located relative to major urban areas within a state. Additionally, PFs tend to enroll more students and employ more staff, meaning they often stabilize local economies to a greater degree than RPUs. The semi-elasticities for return to high school labor market tell a different story: the differences between in-state and out-of-state students are larger, but differences across institutional groupings narrow. While out-of-state graduates of RPUs are slightly less responsive to labor demand shocks in the high school labor market, the variation across institutions largely stems from differences in the share of in-state students in each institutional grouping: according to Table 6, half of Private college alumni attend college in the same

23. Table 7 presents the data in Tabular form
24. Section 6.2.2 explains the measurement of occupational prestige.

state as their high school, while the same measure is 76 percent for PFs and 88 percent for RPUs. In-state and out-of-state students are not randomly assigned: the lower return migration probabilities partially reflect the preferences of out-of-state college graduates to simply live elsewhere (Groen 2004). The current model also does not disentangle the role social networks play. Out-of-state college graduates might have more spatially dispersed social networks, providing a broader set of alternative migration destinations (Koenen and Johnston 2024). When selecting migration destinations, college graduates (particularly graduates of highly selective institutions) tend to select local labor markets in the Untied States based on where their friends reside (Koenen and Johnston 2024; Manduca 2024).

The IV coefficients in Tables 2, 3, and 4 offer a prediction of how the probability of return migration will vary when a labor market experiences a 1 percent labor demand shock. Figure 10 compares the prediction to reality in each labor market, accounting for both the institutional grouping and in-state status of workers from the origin.[25] Even after accounting the mix of in-state status and institutional groupings, some labor markets (like New York City and Chicago) out-perform the IV prediction in Model (5) of Tables 2, 3, and 4. In general, the college graduates from larger cities tend to return at higher rates than the labor demand instrument instrument suggests. The gap between the predicted and actual return rate suggests those who do return compensate for a worse local labor market with amenities. Early works on the migration patterns of college graduates focused on physical amenities, whether "third spaces", walkability, and transit (Florida 2002) or average winter temperature and wind speed (Kodrzycki 2001). More recent work, such as Koenen and Johnston 2024 and Hershbein and Stuart 2024 suggests social ties constitute a type of amenity, since they compensate for a weaker labor market. Furthermore, variation in the size of social networks might explain might explain why a labor market like Detroit out-performs the IV prediction (Koenen and Johnston 2024). For many workers, the prospect of rebuilding decades of social ties proves too daunting, even with the prospect of higher earnings or more job security. Figure 10 does not delineate between social networks and physical amenities because I do not observe the networks of LinkedIn users, only their profiles.

4.2.4 Back of Envelope

Next, I do a back-of-envelope calculation to show how differences in migration semi-elasticities translate into costs to taxpayers. In the era of fiscal competition between government for economic development (Manduca 2019; Pacewicz 2016), state and local governments can theoretically purchase labor demand shocks with the right package of tax incentives.[26] Suppose for an example a local government uses tax incentives to attract

25. Therefore, the spatial variation in the residual plotted in Figure 10 accounts for the fact that more people might attend college at an in-state RPU in Kalamazoo relative to Ann Arbor. The return probability should be higher in Kalamazoo simply because more students attend in-state RPUs, and graduates of those institutions return at higher rates.

26. Of course, many deals fail to create as many jobs as promised, or include unforeseen costs (Bartik 2020).

a new employer, thereby spurring an exogenous shock to demand in the local labor market. The shock to labor demand increases average wages in the labor market (Bartik 1991). According to the model described in Section 8, higher expected wages make the labor market relatively more attractive to people outside the labor market. Yet Figure 8 suggests college graduates' responsiveness to labor demand shocks varies based on their institutional grouping and in-state status. Out-of-state students are less sensitive to changes in the origin labor market, meaning identical population responses require a larger demand shock. I estimate the cost to taxpayers to induce a 1 percent increase in the total population by increasing return migration rates of college graduates. Assume the local labor market is average-sized (280,000 workers), with an incentive cost of a job of $196,000 (Bartik 2020), and all inframarginal returners will be college graduates, with the pool of potential returners reflecting the national composition of the college graduates by institutional group. Regardless of institutional grouping, every type of out-of-state of college graduates requires 10 percent more than the national average cost of $493 million. From the taxpayers' perspective, out-of-state graduates of private colleges are most expensive to recruit, requiring 13 percent more incentives than the national average. Perhaps unsurprisingly, in-state graduates of RPUs are the most economical for taxpayers trying to minimize incentive costs. Recruiting only in-state RPU graduates will lower the cost of a 1 percent population increase by 5 percent. From the perspective of local economic developers, not all college graduates are created equal: some are much cheaper to recruit.

4.3 Robustness Checks

The Appendix includes robustness checks to gauge the durability of the core findings. Section 7.3 preserves the restriction to 2000-13 college graduation cohorts but varies the dependent variable. I include three alternative versions of the dependent variable: return within five years (same as main analysis but the window narrows from 10 years to 5), residence in years 0 and 10 (simple binary indicator for residence in years 0 and 10), and residence in years 0 and 5 (same but comparing years 0 and 5). Sample sizes substantially on the residence variables because they require a non-missing response in years 0 and 5 or 10. By contrast, the dependent variable in the main analysis and Tables 9, 10, 11 allows for some missing data within the years after graduating college.

- Return Within 5 Years: Tables 9, 10, 11
- Residence in Years 0 and 10: Tables 12, 13, 14
- Residence in Years 0 and 5: Tables 15, 16, 17

All three of the alternative dependent variables preserve the significance and direction of the core results in the main analysis. A 1 percent labor demand shocks in the college labor market increase the probability of return by 0.30-0.45 p.p for people who attended high school and college in different labor markets (Tables 9, 12, and 15). A shock of the same magnitude in the high school labor market increases the probability of return to the high school labor market by 0.15-0.25 p.p. (Tables 10, 13, and 16). The return migration response for people who attended high school and college in the same labor market is indistinguishable from zero (Tables 11, 14, 17).

5 Conclusion

Data linking individuals to origins, destinations, and educational institutions remains challenging to obtain for large swathes of the United States. LinkedIn profiles, which include self-reported educational and residential data, present a novel solution. Conzelmann et al. 2023 demonstrated the usability of LI summary data for analysis of sub-state migration patterns in the US. Individual LI profiles act as microdata, opening the door to the analysis of individual decisions, controlling for characteristics salient to migration such as college characteristics (Conzelmann et al. 2023), in-state status (Groen 2004), and major (Bound et al. 2004). Users who list a high school, college, and post-graduation job offer longitudinal data on the spatial distribution of their social ties, far richer than the repeated cross-sections in most public-use microdata. As a result, LI profiles can answer questions about how college graduates' balance local economic conditions and social ties when making migration decisions.

Brain drain is reversible. While college graduates exhibit greater sensitivity to local labor market conditions than other populations, social ties still shape migration decisions (Hershbein and Stuart 2024; Zabek 2024). A one percent positive shock to local labor demand increases the probability of return migration in both the college and high school labor market. College graduates return to labor markets where they likely have friends and family, suggesting social ties modulate location decisions (Koenen and Johnston 2024; Zabek 2024; Winters 2020).

College graduates are not homogeneous; migration responses vary depending on where someone obtained their bachelor's degree (Conzelmann et al. 2023). In general, graduates of in-state public institutions respond to improvements in local labor market conditions more strongly than peers from private institutions. Similarly, transfer status and age at graduation shape the relative weight graduates place on the high school and college labor market respectively. The variation in responsiveness to shocks in origin labor markets means some populations generate more graduates who live in-state per dollar of state appropriations. Future research should explore the return on investment to state appropriations across different institutional groups, accounting for the variation in migration outcomes.

The results pertain to public policy discussions regarding the funding of post-secondary education systems (Trostel 2010; Strathman 1994). Policymakers increasingly frame higher education as a form of economic development, so return on investment for institutions matters more than ever. Yet RPUs face immense challenges in the coming decades as state support faces increasing competition and student populations continue to decline (Kelchen 2023; Bound et al. 2019). RPUs and Public Flagships respond to tightening state appropriations in distinct ways: flagships often make up for shortfalls by recruiting out-of-state and international students while RPUs struggle to increase enrollment (Bound et al. 2019; Bound, Hershbein, and Long 2009). Instead, RPUs cut instructional costs and services to students (Bound et al. 2019). Institutions most vulnerable to the economic and demographic vicissitudes of the next few decades produce graduates with the strongest attachment to their home labor market and state.

References

Abadie, Alberto. 2003. "Semiparametric instrumental variable estimation of treatment response models." *Journal of Econometrics.*

Acton, Riley, Kalena Cortes, Lois Miller, and Camila Morales. 2024. *Distance to Degrees: How College Proximity Shapes Students' Enrollment Choices and Attainment Across Race-Ethnicity and Socioeconomic Status.* Working Paper, Working Paper Series 33085. National Bureau of Economic Research.

Alm, James, and John V. Winters. 2009. "Distance and intrastate college student migration." *Economics of Education Review* 28 (6).

Amior, Michael, and Alan Manning. 2018. "The Persistence of Local Joblessness." *American Economic Review.*

Angrist, Joshua, Guido Imbens, and Alan Krueger. 1999. "Jacknnife Instrumental Variables Estimation." *Journal of Applied Econometrics.*

Autor, David, and Mark Duggan. 2003. "The Rise in the Disability Rolls and the Decline in Unemployment." *The Quarterly Journal of Economics.*

Bartik, Timothy. 2020. "Using Place-Based Jobs Policies to Help Distressed Communities." *Journal of Economic Perspectives.*

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute.

Berry, Alexander, Elizabeth M. Maloney, and David Neumark. 2024. "The Missing Link? Using LinkedIn data to measure race, ethnic, and gender differences in employment outcomes at individual companies." *National Bureau of Economic Research.*

Blanchard, Olivier, and Larry Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity.*

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2025. "A Pratical Guide to Shift-Share Instruments." *Journal of Economic Perspectives.*

———. 2022. "Quasi-Experimental Shift-Share Research Designs." *The Review of Economic Studies.*

Bound, John, Breno Braga, Gaurav Khanna, and Sarah Turner. 2019. *Public Universities: The Supply Side of Building a Skilled Workforce.* Technical report. National Bureau of Economic Research.

Bound, John, Jeffrey Groen, Gábor Kézdi, and Sarah Turner. 2004. "Trade in university training: cross-state variation in the production and stock of college-educated labor." *Journal of Econometrics.*

Bound, John, Brad Hershbein, and Bridget Terry Long. 2009. "Playing the Admissions Game: Student Reactions to Increasing College Competition." *The Journal of Economic Perspectives* 23 (4): 119–146.

Bound, John, and Harry J Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics* 18 (1): 20–54.

Bárány, Zsófia, Moshe Buchinsky, and Pauline Corblet. 2023. *Late Bloomers: the Aggregate Implications of Getting Education Later in Life.* Working Paper, Working Paper Series 31874. National Bureau of Economic Research.

Cadena, Brian C., and Brian K. Kovak. 2016. "Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession." *American Economic Journal: Applied Economics* 8 (1): 257–90.

Cajner, Tomaz, John Coglianese, and Joshua Montes. 2021. *The Long-Lived Cyclicality of the Labor Force Participation Rate.* Technical report. Federal Reserve Bank Board of Governors Finance and Discussion Series.

Chan, Sewin, Katherine O'Regan, and Wei You. 2020. "Migration choices of the boomerang generation: Does returning home dampen labor market adjustment." *Journal of Housing Economics.*

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2020. "Income Segregation and Intergenerational Mobility Across Colleges in the United States*." *The Quarterly Journal of Economics* 135, no. 3 (February): 1567–1633.

———. 2017. *Mobility Report Cards: The Role of Colleges in Intergenerational Mobility.* Technical report.

Condon, David, and Bradley Hughes. 2022. *Occupational Prestige Ratings Data.* V. V2.

Conzelmann, Johnathan G, Steven W Hemelt, Brad Hershbein, Shawn M Martin, Andrew Simon, and Kevin M Stange. 2022. *Grads on the Go: Measuring College-Specific Labor Markets for Graduates.* Working Paper, Working Paper Series 30088. National Bureau of Economic Research.

———. 2023. "Grads on the Go: Measuring College-Specific Labor Markets for Graduates." *Journal of Policy Analysis and Management,* 1–22.

Curtis, Mark, Layla O'Kane, and R. Jisung Park. 2024. "Workers and the Green Energy Transition: Evidence from 300 Million Job Transitions." *Environmental and Energy Policy and the Economy.*

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review* 106 (3): 479–524.

Diamond, Rebecca, and Cecile Gaubert. 2022. "Spatial Sorting and Inequality." *Annual Review of Economics* 14 (1): 795–819.

Diamond, Rebecca, and Enrico Moretti. 2021. *Where is Standard of Living the Highest? Local Prices and the Geography of Consumption.*

Dorling, Daniel. 1996. *Area Cartograms: Their Use and Creation.* Technical report, Concepts and Techniques in Modern Geography. University of East Anglia.

Dorn, David, Florian Schoner, Moritz Seebacher, Lisa Simon, and Ludger Woessmann. 2025. *Multidimensional Skills as a Measure of Human Capital: Evidence from LinkedIn Profiles.* Technical report. ArXiv.

Eckert, Fabian, Teresa C. Fort, Peter K. Schott, and Natalie J. Yang. 2021. *Imputing Missing Values in the US Census Bureau's County Business Patterns.* Technical report. National Bureau of Economic Research.

Evsyukova, Yulia, Felix Rusche, and Wladislaw Mill. 2025. "LinkedOut? A Field Experiment on Discrimination in Job Network Formation." *Quarterly Journal of Economics.*

Florida, Richard. 2002. *The Rise of the Creative Class.* Basic Books.

Gaubert, Cecile, Patrick Kline, Damián Vergara, and Danny Yagan. 2025. "Place-Based Redistribution." *American Economic Review.*

Glaeser, Edward, and Joshua Gottlieb. 2008. "The Economics of Place-Making Policies." *Brookings Paper on Economic Activities.*

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review* 110 (8): 2586–2624.

Gortmaker, Jeff, Jessica Jeffers, and Michael Junho Lee. 2024. *Labor Reactions to Credit Deterioration: Evidence from LinkedIn Activity.* Technical report.

Groen, Jeffrey A. 2004. *The Effect of College Location on Migration of College-Educated Labor.* Technical report. University of Michigan Population Studies Center.

Guillory, Jamie, and Jeffrey T. Hancock. 2012. "The effect of Linkedin on deception in resumes." *Cyberpsychology, Behavior, and Social Networking.*

Hershbein, Brad, and Bryan Stuart. 2024. "The Evolution of Local Labor Markets After Recessions." *American Economic Journal: Applied Economics.*

Hillier, Cathlene, Yujiro Sano, Roger Pizzarro Milian, and David Zarifa. 2023. "Where Did They Go? Regional Patterns in Early Transfer in Ontario Post-Secondary Education." *Canadian Journal of Higher Education.*

Kelchen, Robert. 2023. "The Haves and Have Nots of Higher Education: College Enrollments are Becoming Increasingly Unequal." *The Chronicle of Higher Education.*

Kim, Bokyung, Minseog Kim, and Geunyong Park. 2025. *Pharmacy School Cost and Opioid Dispensing: A Hidden Connection?* Technical report.

Kline, Patrick, and Enrico Moretti. 2014. "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *Quarterly Journal of Economics.*

Kodrzycki, Yolanda. 2001. "Migration of Recent College Graduates: Evidence form the National Longitudinal Survey of Youth." *New England Economic Review.*

Koenen, Martin, and Drew Johnston. 2024. "Social Ties and Residential Choice: Micro Evidence and Equilibrium Implications." PhD diss., Harvard University.

Kreisman, Daniel, Jonathan Smith, and Bondi Arifin. 2021. *Labor Market Signaling and the Value of College: Evidence from Resumes and the Truth.* IZA Discussion Papers 14483. Institute of Labor Economics (IZA).

Krolikowski, Pawel, Mike Zabek, and Patrick Coate. 2020. "Parental Proximity and Earnings After Job Displacements." *Labour Economices.*

Manduca, Robert A. 2024. "I Was Open to Anywhere, It's Just This Was Easier: Social Structure, Location Preferences, and the Geographic Concentration of Elite College Graduates." *Qualitative Sociology.*

Manduca, Robert A. 2019. "The Contribution of National Income Inequality to Regional Economic Divergence." *Social Forces* 98 (2): 622 –648.

Moretti, Enrico. 2003. "Human Capital Externalities in Cities," NBER working paper series; no. w9641.

———. 2012. *The New Geography of Jobs.* Houghton Mifflin Harcourt.

Mountjoy, Jack. 2024. *Marginal Returns to Public Universities.* Technical report. University of Chicago.

Notowidigdo, Matthew J. 2020. "The Incidence of Local Labor Demand Shocks." *Journal of Labor Economics* 38 (3): 687–725.

Pacewicz, Josh. 2016. *Partisans and Partners: The Politics of the Post-Keynesian Society.* University of Chicago Press.

Piyampromdee, Suphanit. 2021. "The impact of immigration on wages, internal migration, and welfare." *The Review of Economic Studies.*

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (6): 1257–1278.

Sprung-Keyser, Ben, Nathaniel Hendren, and Sonya Porter. 2022. *The Radius of Economic Opportunity: Evidence from Migration and Local Labor Markets.* Technical report.

Strathman, James G. 1994. "Migration, Benefit Spillovers and State Support for Higher Education." *Urban Studies.*

Trostel, Philip A. 2010. "The Fiscal Impacts of College Attainment." *Research in Higher Education.*

Winters, John V. 2020. "In-State College Enrollment and Later Life Location Decisions." *Journal of Human Resources.*

Wozniak, Abigail. 2010. "Are College Graduates More Responsive to Distant Labor Market Opportunities?" *Journal of Human Resources* 45 (4): 944–970.

Yagan, Danny. 2019. "Employment Hysteresis from the Great Recession." *Journal of Political Economy.*

Yagan, Danny. 2014. *Moving to Opportunity?: Migratory Insurance Over the Great Recession.* Technical report.

Zabek, Mike. 2024. "Local Ties in Spatial Equilibrium." *American Economic Journal: Macroeconomics.*

Table 1: Summary statistics of sample relative to universe

	All Col. Graduates	All HS Includers	All Grads in Cohort	Primary Sample
Source	ACS	LI	ACS	LI
Col. Graduation Cohorts	All	All	2000-13	2000-13
Public Flagship	24	23	23	22
Private, For Profit	2	0	2	0
Private, Non-Profit	32	41	32	42
Regional Public	43	36	43	36
Mobility Rate	2	2	2	2
Mean Parent Income ($1,000)	143.2	167.8	142.1	168.9
Mean Kid Income ($1,000)	57.0	62.7	56.6	62.0
Average Acceptance Rate	66	63	66	63
90th Percentile	1959	1969	1975	1975
75th Percentile	1968	1980	1979	1980
Median	1979	1991	1983	1985
25th Percentile	1989	1996	1987	1989
10th Percentile	1995	1999	1989	1991
Northeast	22	27	22	24
South	32	27	33	29
Midwest	25	25	25	26
West	20	19	21	19
N	67,217,457	2,476,759	23,199,063	474,250

Summary statistics of the sample relative to alternative samples of college-educated workers. Rows with "ACS" listed as source use 2023 American Community Survey 1-year estimates. The first and third columns report ACS data, restricting to labor force participants with at least a bachelor's degree. The third column additionally restricts to the 2000-13 college graduation cohorts. The second column characterizes all LI users listing both a high school and college, while the fourth column reports characteristics of LI users in the final sample: workers who completed college between 2000 and 2013 and left their origin labor market at some point after college. Section 6.1 describes the construction of the sample in greater detail. The ACS lacks questions about where people obtained their bachelor's degree, so I simulate the distribution of baccalaureate institutions among US workers using institution-level degree awards from 1983-2022 in IPEDS. For pre-1983 graduation cohorts (8 percent of workers in 2023), I used the institutional distribution of degrees awarded in 1983. The second to sixth rows reports the simulated shares of graduates by institutional grouping. Rows 7-10 estimate institution-level summary statistics using shares derived from the simulated distribution and values from Chetty et al. 2017. Note that I computed the values by reweighting institution-level aggregate statistics, rather than directly observing income of workers. Rows 11-15 report birth cohort distributions. In the LI data, I estimate year of birth and observe year of college graduation. In the ACS data, I observe year of birth and estimate year of college graduation using the joint distribution created from LI data. The bottom four rows report the share of bachelor's degrees awarded in each Census region, based on the location of the college.

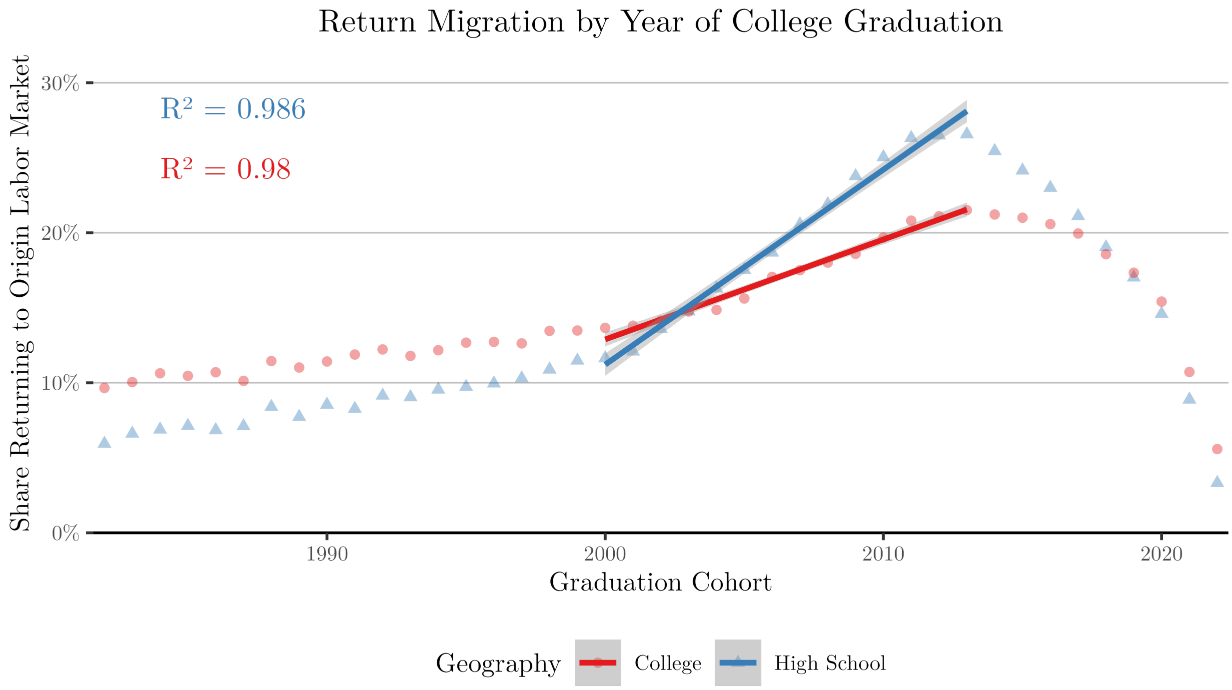

Figure 1: Share of workers returning to the origin labor market within 10 years of college graduation, by college graduation cohort. Alternatively, the average value of the binary dependent variable in Model 6. The red points measure the dependent variable in Table 2, while the blue measures the dependent variable in Tables 3 and 4. While the main analysis in Section 4 only includes college graduation cohorts between 2000 and 2013, the figure includes cohorts from 1982 to 2022. Lines of best fit for each return geography weight each year between 2000 and 2013 equally, with R^2 reported on the plot.

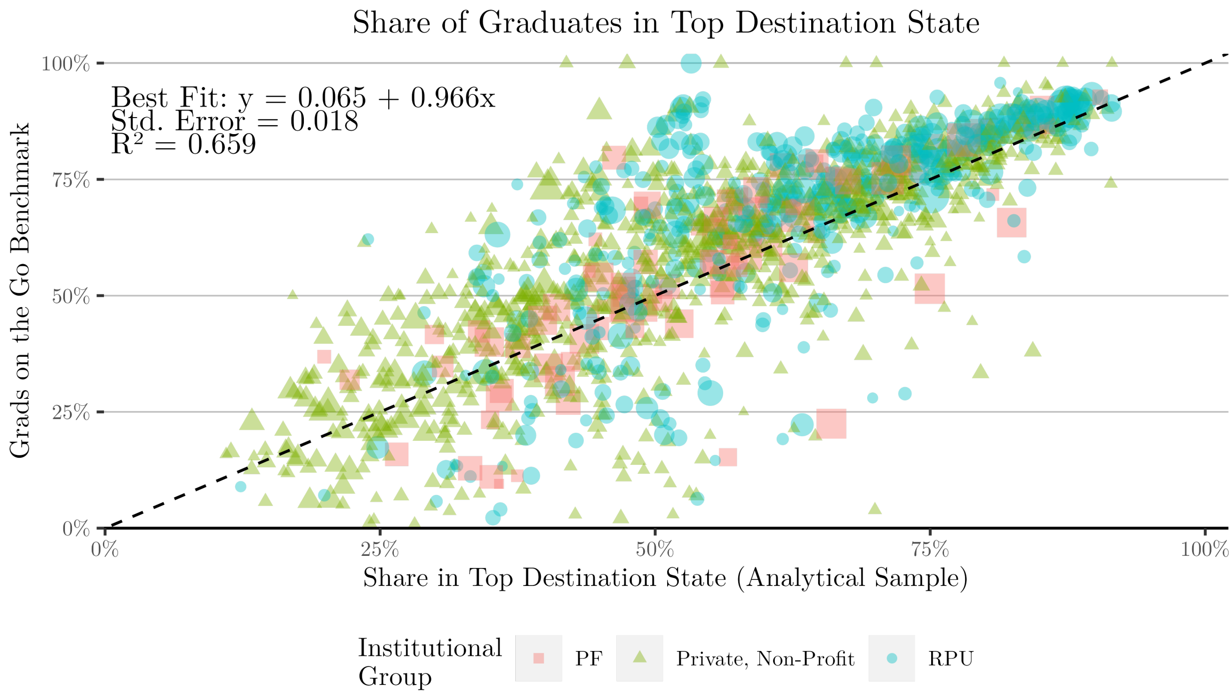

Figure 2: The analytical sample shows similar spatial patterns to Conzelmann et al. 2023, even though they used summary data from institutional LI pages rather than microdata. Every circle represents a college or university, sized by number of degrees awarded between 2000 and 2022. The dashed line is $y = x$, with the best fit line, R^2, and standard errors weighted by degrees awarded. Broadly speaking, the top destination state captures a similar share of respondents to Conzelmann et al. 2023. A 1 p.p. increase the share of graduates destined for the top state in Conzelmann et al. 2023 is associated with a 0.966 p.p. increase in the analytical sample.

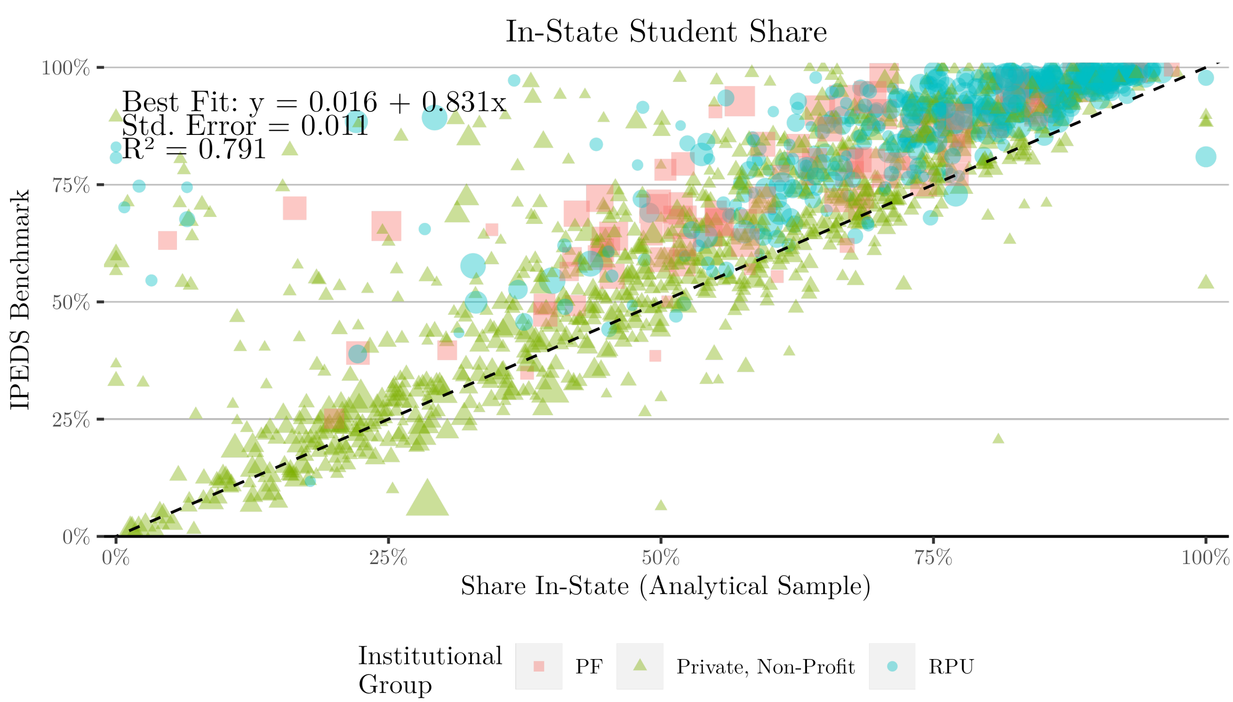

Figure 3: The analytical sample reflects the spatial patterns reported in IPEDS institution-level data on in-state students. Every circle represents a college or university, sized by number of degrees awarded between 2000 and 2022. The dashed line is $y = x$, with the best fit line, R^2, and standard errors weighted by degrees awarded between 2000 and 2022. In-state student shares reported by IPEDS accord with the results in the analytical sample, with a lower standard error than post-graduation migration outcomes in Figure 2. A 1 p.p. increase in an institution's average in-state share is associated with a 0.831 p.p. increase in the in-state share in the analytical sample.

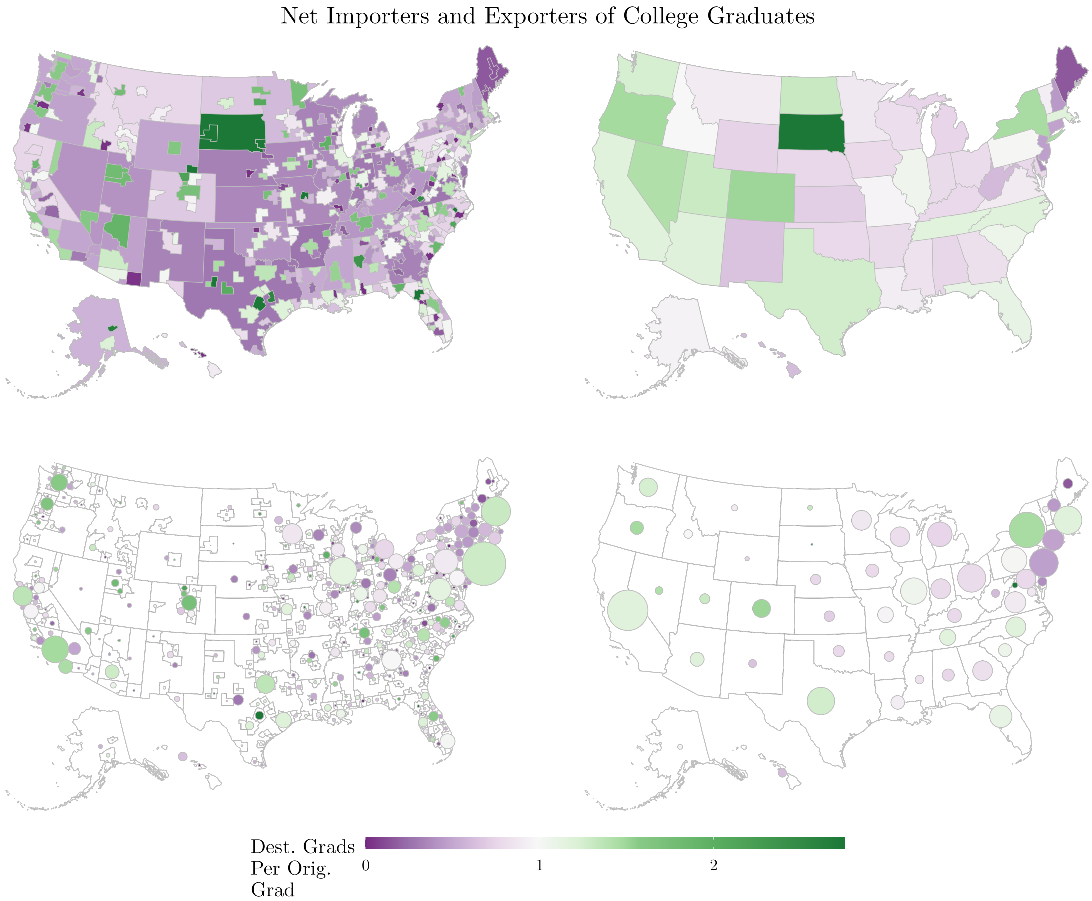

Figure 4: The maps show how net flows of college graduates vary by state and labor market. Net flow is the ratio of observations working in the labor market after graduating college to the number graduating from high school in a location. The top row shows the traditional chloropleth map, while the bottom presents a Dorling cartogram weighted by the number of graduates originating in the labor market Dorling 1996. The net flow measure is agnostic to whether graduates remain in their origin destination, focusing on aggregate flows of graduates instead. Attracting more graduates than originate in the location results in a ratio greater than 1. Thus, labor markets colored green are net receipients of college gaduates in the LI sample, while purple ones donate college graduates. Gray locations attract approximately as many college graduates as they produce. To limit the visual impact of outliers with small sample sizes, I applied Empirical Bayes shrinkage on the net migration flows. In the analytical sample, almost every local labor market produces a non-zero number of college-bound high school graduates. In general, the maps suggest migration flows on LinkedIn replicate the skill sorting observed across US Census data over the last four decades (Diamond 2016; Diamond and Gaubert 2022; Diamond and Moretti 2021). Large well-educated urban areas net college graduates. Smaller urban areas and rural labor markets without major universities export college graduates.

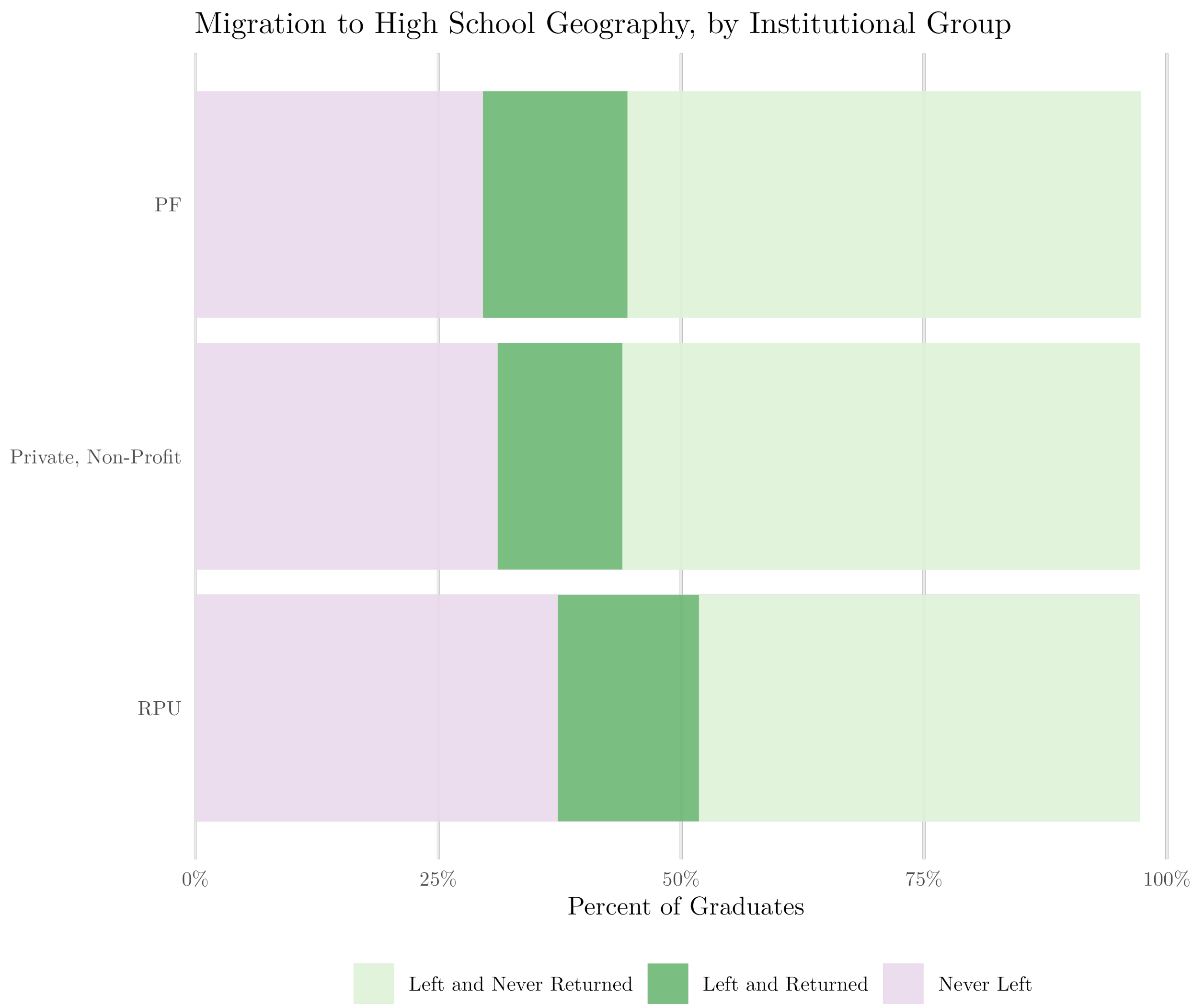

Figure 5: Post-graduation migration decisions vary considerably based on institutional grouping. The three components of each horizontal bar represent mutually exclusive and collectively exhaustive outcomes for college graduates. Individuals encompassed by the lavender bar never live anywhere besides their origin labor market. Individuals in both light and dark green segments lived somewhere besides their origin labor market. The darker component represents individuals who returned to their origin labor market. The dependent variable for the regression coefficients in Tables 2, 3, and 4 is the dark green share relative to the sum of dark and light green.

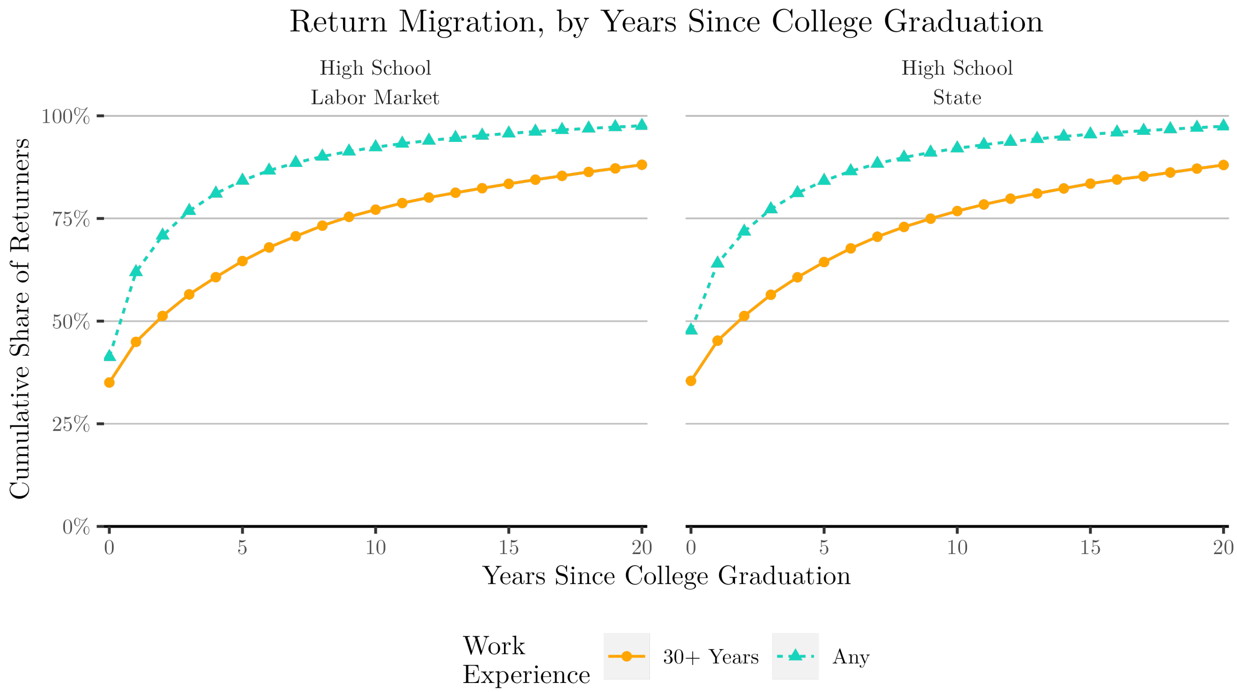

Figure 6: If a college graduate returns to their origin labor market or state, they typically move within the first five years after graduating college. The teal line shows the entire sample, without restrictions. Given the sample's bias towards younger workers (see Figures 15 and 16), I include workers who include at least 30 years of work history too. While fewer workers with extensive work history return at every year, more than 75 percent of return migration occurs within the eight years following graduation. Migration decisions for college graduates largely occur in a narrow window following the completion of a bachelor's degree.

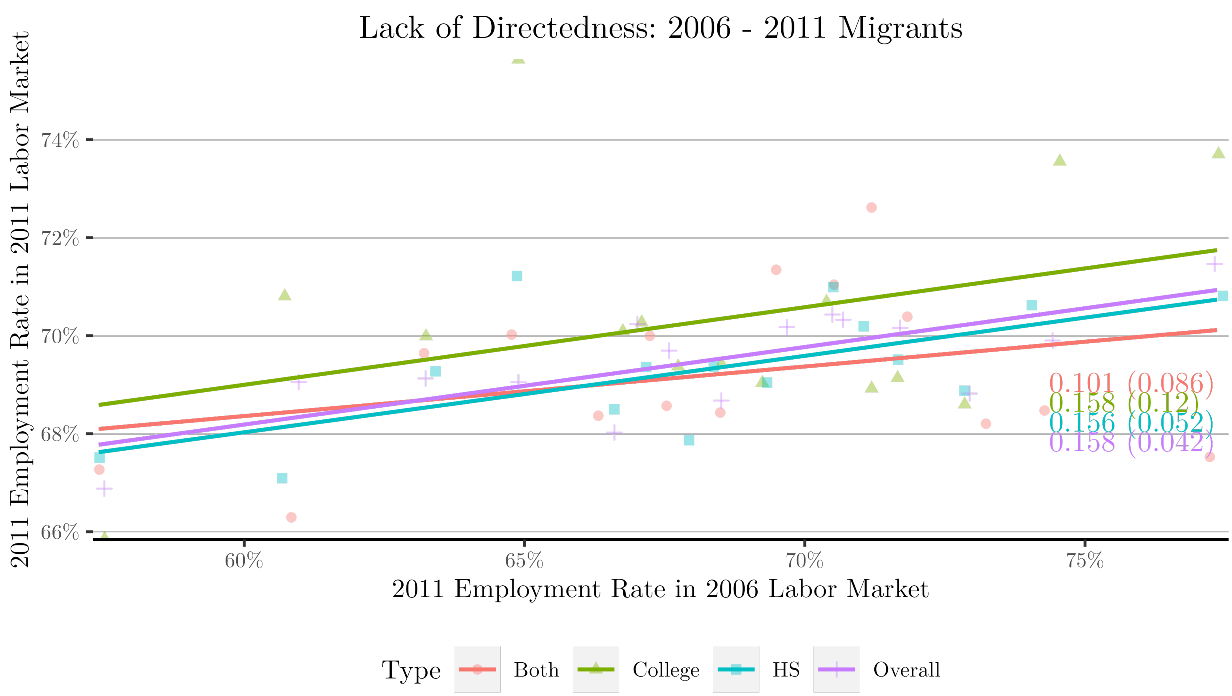

Figure 7: Return migration in LI data exhibits the same lack of directedness observed by Yagan 2014, who finds a slope of 0.195 with IRS administrative data. If workers move labor markets between 2006 and 2011, how do the 2011 employment rates compare between the origin (x-axis) and destination (y-axis)? Each line measures the line of best fit for a return migration channel based on the individual's connection to the labor market. If the equilibrating migration predicted by Roback 1982 occurred, the line would have a slope of -1. A positive slope coefficient suggests migration is not directed, meaning migrants to high-employment labor markets often originate in similar ones (Yagan 2014).

Table 2: Probability of returning to labor market containing college within 10 years of college graduation

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	0.001	−0.030	−0.026	−0.022	-0.065^{*}	-0.065^{*}	−0.036	−0.031	−0.025	-0.077^{**}
	(0.030)	(0.030)	(0.031)	(0.030)	(0.027)	(0.033)	(0.033)	(0.033)	(0.032)	(0.029)
College LD Shock	0.397^{***}	0.381^{***}	0.388^{***}	0.367^{***}	0.339^{***}	0.424^{***}	0.459^{***}	0.463^{***}	0.438^{***}	0.405^{***}
	(0.023)	(0.028)	(0.028)	(0.026)	(0.025)	(0.026)	(0.030)	(0.031)	(0.029)	(0.027)
Private			0.077^{***}	0.076^{***}	0.080^{***}			0.077^{***}	0.076^{***}	0.080^{***}
			(0.007)	(0.006)	(0.005)			(0.007)	(0.006)	(0.005)
RPU			0.125^{***}	0.124^{***}	0.126^{***}			0.125^{***}	0.124^{***}	0.126^{***}
			(0.007)	(0.006)	(0.006)			(0.007)	(0.006)	(0.006)
Private × In-State			−0.010	−0.022	−0.015			−0.010	−0.021	−0.015
			(0.014)	(0.014)	(0.012)			(0.014)	(0.014)	(0.012)
RPU × In-State			-0.075^{***}	-0.077^{***}	-0.073^{***}			-0.075^{***}	-0.077^{***}	-0.073^{***}
			(0.014)	(0.014)	(0.013)			(0.014)	(0.014)	(0.013)
College In State			0.046^{***}	0.061^{***}	0.054^{***}			0.046^{***}	0.061^{***}	0.054^{***}
			(0.012)	(0.012)	(0.012)			(0.012)	(0.012)	(0.012)
BA by Age 24				-0.077^{***}	-0.076^{***}				-0.077^{***}	-0.076^{***}
				(0.005)	(0.004)				(0.005)	(0.004)
Transfer Student				0.022^{***}	0.022^{***}				0.022^{***}	0.022^{***}
				(0.004)	(0.004)				(0.004)	(0.004)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
Num. obs.	396546	396546	396546	396546	396546	396546	396546	396546	396546	396546

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

First-stage partial F-statistic for high school LD shock is 4,795,502 ($R^2 = 0.055$) and 6,441,060 for college shock ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor markets as endogenous independent variables, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) incorporates college graduation cohort fixed effects, Model (3) controls for institutional grouping and residency, and Model (4) adds major, transfer, and on-time completion controls. Model (5) estimates Equation 6. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate ever returned to the labor market containing their **college** between the first and tenth years after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to college labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. Table 8 reports results of the first-stage estimation process described in Section 4.1.

Table 3: Probability of returning to labor market containing high school within 10 years of college graduation

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	0.350***	0.208***	0.207***	0.202***	0.190***	0.211***	0.248***	0.250***	0.245***	0.213***
	(0.060)	(0.040)	(0.038)	(0.038)	(0.036)	(0.047)	(0.047)	(0.047)	(0.046)	(0.039)
College LD Shock	−0.014	−0.122***	−0.120***	−0.103***	−0.100***	−0.185***	−0.143***	−0.133***	−0.113***	−0.113***
	(0.030)	(0.035)	(0.030)	(0.029)	(0.019)	(0.033)	(0.036)	(0.031)	(0.030)	(0.020)
Private			0.001	−0.000	−0.002			0.001	−0.000	−0.002
			(0.009)	(0.009)	(0.004)			(0.009)	(0.009)	(0.004)
RPU			−0.034***	−0.032***	−0.025***			−0.033***	−0.032***	−0.025***
			(0.009)	(0.008)	(0.006)			(0.009)	(0.008)	(0.006)
Private × In-State			−0.057***	−0.049***	−0.060***			−0.057***	−0.049***	−0.060***
			(0.011)	(0.011)	(0.008)			(0.011)	(0.011)	(0.008)
RPU × In-State			0.026	0.024	0.013			0.025	0.024	0.013
			(0.014)	(0.014)	(0.009)			(0.014)	(0.013)	(0.009)
College In State			0.108***	0.095***	0.109***			0.108***	0.095***	0.109***
			(0.015)	(0.015)	(0.009)			(0.014)	(0.015)	(0.009)
BA by Age 24				0.078***	0.074***				0.078***	0.074***
				(0.008)	(0.007)				(0.008)	(0.007)
Transfer Student				−0.000	−0.001				−0.000	−0.001
				(0.002)	(0.003)				(0.002)	(0.003)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
Num. obs.	396546	396546	396546	396546	396546	396546	396546	396546	396546	396546

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

First-stage partial F-statistic for high school LD shock is 4,795,502 ($R^2 = 0.055$) and 6,441,060 for college shock ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor markets as endogenous independent variables, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) incorporates college graduation cohort fixed effects, Model (3) controls for institutional grouping and residency, and Model (4) adds major, transfer, and on-time completion controls. Model (5) estimates Equation 6. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate ever returned to the labor market containing their **high school** between the first and tenth years after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to college labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. Table 8 reports results of the first-stage estimation process described in Section 4.1.

Table 4: Probability of returning to labor market containing both high school and college within 10 years of college graduation

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	0.655***	0.132	0.139	0.130	0.109	0.079	0.193	0.201*	0.191	0.129
	(0.142)	(0.087)	(0.086)	(0.087)	(0.079)	(0.087)	(0.099)	(0.098)	(0.099)	(0.085)
Private			−0.018	−0.016	−0.033			−0.018	−0.015	−0.033
			(0.053)	(0.052)	(0.043)			(0.053)	(0.052)	(0.043)
RPU			0.016	0.024	0.051			0.017	0.024	0.051
			(0.062)	(0.061)	(0.056)			(0.062)	(0.061)	(0.056)
Private × In-State			0.039	0.034	0.044			0.039	0.034	0.044
			(0.057)	(0.056)	(0.045)			(0.057)	(0.056)	(0.045)
RPU × In-State			0.003	−0.002	−0.025			0.003	−0.002	−0.025
			(0.064)	(0.063)	(0.055)			(0.064)	(0.063)	(0.055)
College In State			0.010	0.016	0.028			0.010	0.016	0.028
			(0.062)	(0.062)	(0.057)			(0.062)	(0.062)	(0.057)
BA by Age 24				0.022**	0.023***				0.022**	0.023***
				(0.007)	(0.007)				(0.007)	(0.007)
Transfer Student				0.029***	0.024***				0.029***	0.024***
				(0.005)	(0.005)				(0.005)	(0.005)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46
Num. obs.	77695	77695	77695	77695	77695	77695	77695	77695	77695	77695

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

First-stage F-statistic for high school and college LD shock is 1,591,231 ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor market as an endogenous independent variable, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) controls for institutional grouping and residency, Model (3) adds major, transfer, and on-time completion controls, and Model (4) incorporates college graduation cohort fixed effects with a linear time trend. Model (5) estimates Equation 6, just with one independent variable instead of two. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate ever returned to the labor market containing their **high school and college** between the first and tenth years after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to high school and college labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. First-stage partial F-statistic for the high school and college labor demand shock is ($R^2 = 0.835$). Table 8 reports results of the first-stage estimation process described in Section 4.1.

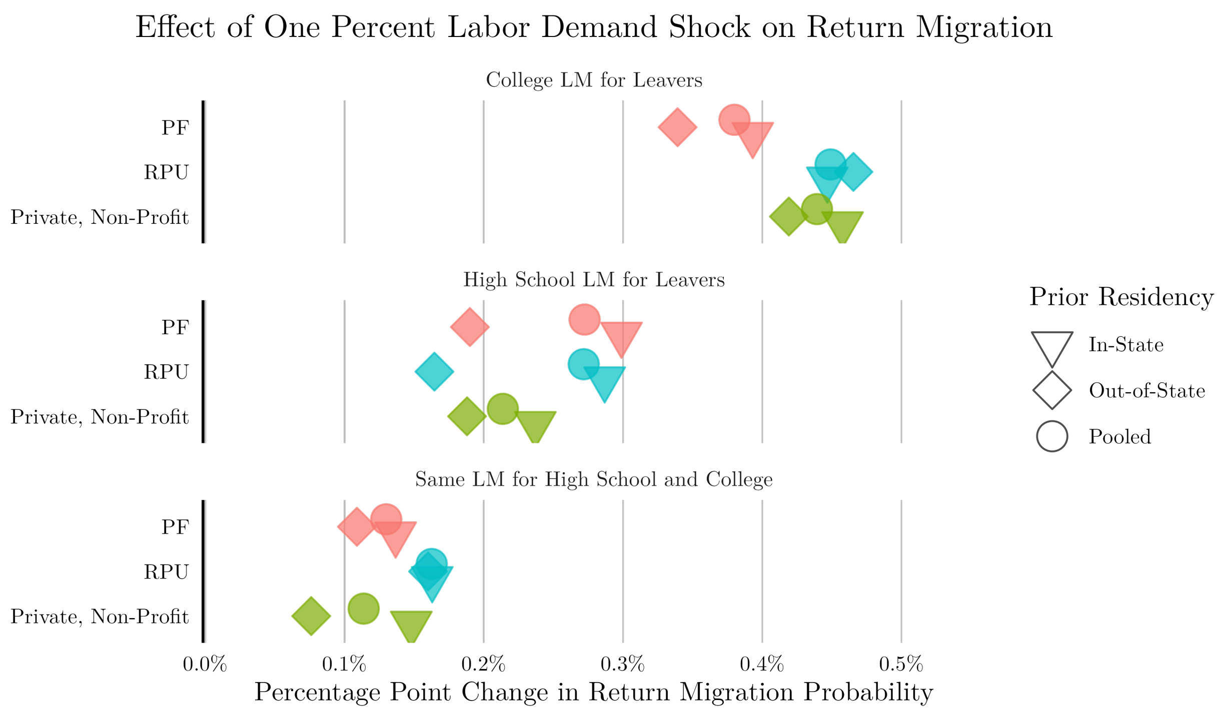

Figure 8: The x-axis plots implied semi-elasticity coefficients from Tables 2, 3, and 4, with each shape representing a different residency status and color signifying an institutional grouping. The pooled measure weights in-state and out-of-state students based on the in-state shares in Table 1. If a worker graduated college in a different labor market than high school and moved away from both after college, the unconditional mean probabilities of return to high school and college labor markets are 0.33 and 0.17 respectively. If the worker completed high school and college in the same labor market, the unconditional mean is 0.46.

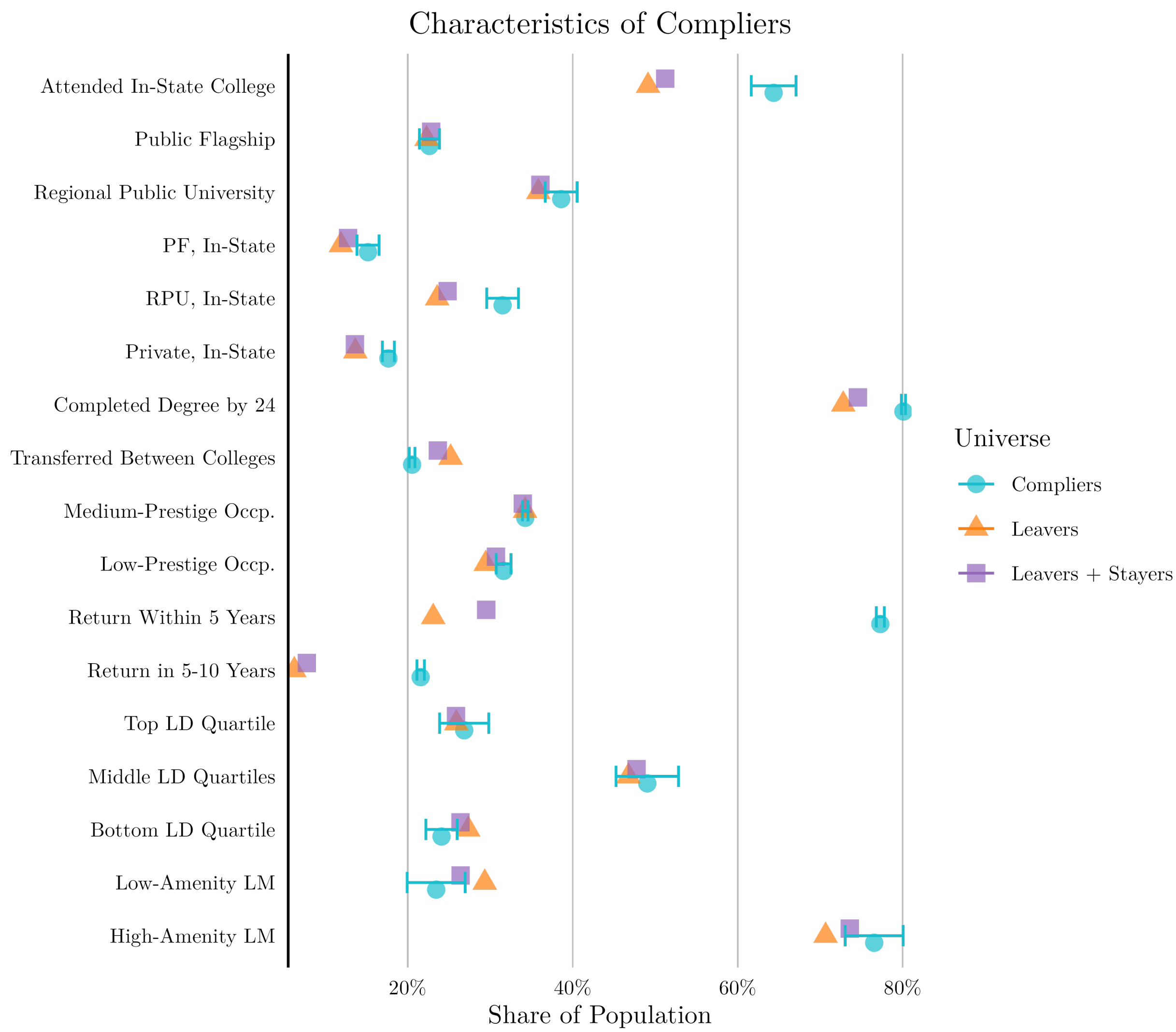

Figure 9: Compliers are college graduates induced to return to their high school labor market by labor demand shocks. Estimation of complier means relies on the Kappa theorem in Abadie 2003. The blue elements with error bars are the complier means and their White standard errors. The orange triangles represent shares for includes all college graduates who moved at some point after graduating college between 2000 and 2013. The purple square includes anybody graduating between 2000 and 2013, regardless of whether they moved after college. Compliers tend to graduate from Regional Public Universities and baccalaureate institutions in the same state as their high schools. In-state graduates are nearly two-thirds of compliers, despite accounting for half of the sample. College graduates who return to origin labor markets mirror the occupational composition of the labor market, with each occupational prestige bin accounting for about one-third of compliers. Table 7 presents the data in tabular form.

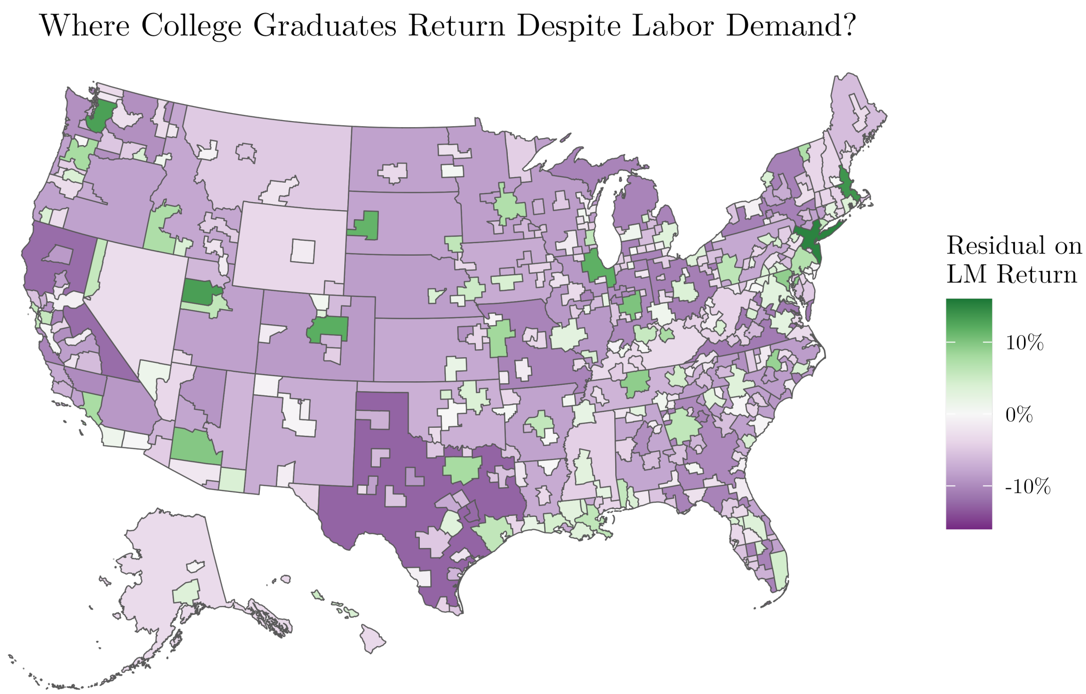

Figure 10: The migration semi-elasticities in Tables 2, 3, and 4 and plotted in Figure 8 are predictions of return migration based on labor demand shocks. Yet substantial heterogeneity exists within American labor markets, with some labor markets enjoying higher rates of return migration than expected. To limit the visual impact of outliers with small sample sizes, I applied Empirical Bayes shrinkage on the residual. The largest labor market in each state tends to recruit more college graduates than expected given its labor demand. The over-performance implicitly measures amenities as a bundle of goods. Amenities might include nice weather, access to transit, or other physical characteristics (Kodrzycki 2001; Florida 2002). Koenen and Johnston 2024 demonstrates the importance of average social network size, which varies across labor markets. Labor markets where workers tend to enjoy larger labor markets, like Detroit, exceed the model's predictions because larger social networks insure against local shocks to employment (Koenen and Johnston 2024; Zabek 2024).

6 LinkedIn Data Appendix

6.1 Dataset Construction

6.1.1 Structure of Data

I created a database of individual-level educational attainment and migration behavior using LinkedIn profile data. Revelio Labs developed the data, then the Brookings Institution's Workforce of the Future Research Initiative purchased access to the data.

The raw data from Revelio consists of an education .csv file and two position .csv files totaling 60 gigabytes. I received the data on 17 July 2023. In the education file, each row is an educational experience attached to a user ID. For example, a user listing a bachelor's degree and a master's degree would appear as two observations, one for each degree, even if they received the degree from the same institution. The education file also includes fields for start date, end date, degree awarded, and major. The major is imputed by Revelio. Similarly, the position files consist of jobs linked to user IDs. On average, users reported five positions. Even if the positions were at the same firm, they appear as multiple observations. Importantly, the position file includes fields for occupation title, occupational SOC code (imputed by Revelio and basis for occupational prestige rating), location, start date, and end date.

Table 5: Observation accounting

Reason	Sample
All BA Completers in US	44114167
Include HS	3032586
Work history has geodata	2939156
Work history has end date after start	2912279
Work history has start or end date	2912128
Work history between 1975 and 2023	2575238
Born [1930,2002]	2476759
Left origin labor markets after graduating from college [2000,2013]	474250

Requiring college graduates to include a high school is the single costliest requirement for the size of the analytical sample. The following sections describe the reasoning behind the sample restrictions.

6.1.2 Construction Steps

The raw data does not include standardized codes for educational institutions. In many cases, the degree field is empty too. Perhaps unsurprisingly, listing a high school on a LinkedIn profile isn't typical, so the data cleaning process revolves around identifying plausible high school names. Consequently, the linkage of users to high school and college unfolds as follows:

1. Constructing crosswalks between LinkedIn geographies on labor markets

   (a) The geographic data in the raw forms is not so pleasant. Instead of a county or CBSA FIPS, a string fills the geography field. The string report a name (which almost always matches an MSA). Notably, the string also includes a single state, even in multi-state metropolitan areas. The state denotes the location of whatever the user reports, rather than the largest location of the core city in the metropolitan area. Note that Metropolitan Divisions are not reported in the data.

   (b) To simplify cleaning, I take the principal city in the metropolitan area, removing any names that appear after a hyphen. Over time, the secondary cities in metropolitan areas change, though principal cities rarely do.

   (c) I correct for (rare) instances of misspelled city names or changes in the principal city in a metropolitan area:

      - Prescott → Prescott Valley
      - Santa Barbara → Santa Maria
      - Fort Walton Beach → Crestview
      - Sarasota → North Port
      - Honolulu → Urban Honolulu
      - Corvalis → Corvallis

   (d) While I now have a cleaned principal city and state, FIPS codes are preferable for predictable merges. Ultimately, I want every observation attached to defined geographic areas, rather than nebulous strings. So I merge the string names of primary cities in 2023 metropolitan areas on a crosswalk to county FIPS codes.[27] Obviously, non-metropolitan areas exist across the United States. I group all non-metropolitan counties by state, forming groups which may or may not exhibit contiguity. Figure 11 shows the final sub-state labor markets across the United States.

2. Identifying users with both high school and college listed:

   (a) Remove educational experiences with "Associate" in the degree field.

   (b) Remove educational experiences with a country besides the United States listed. The identification strategy requires exposure to high school and college labor markets in the United States. The inclusion of immigrants would likely attenuate coefficients even more, given the sensitivity of foreign-born workers to labor demand shocks Cadena and Kovak 2016; Piyampromdee 2021

27. The only exception is the State of Connecticut, where a 2022 reorganization of county-equivalents would have complicated analysis. The raw Connecticut data uses the old boundaries, so I do too.

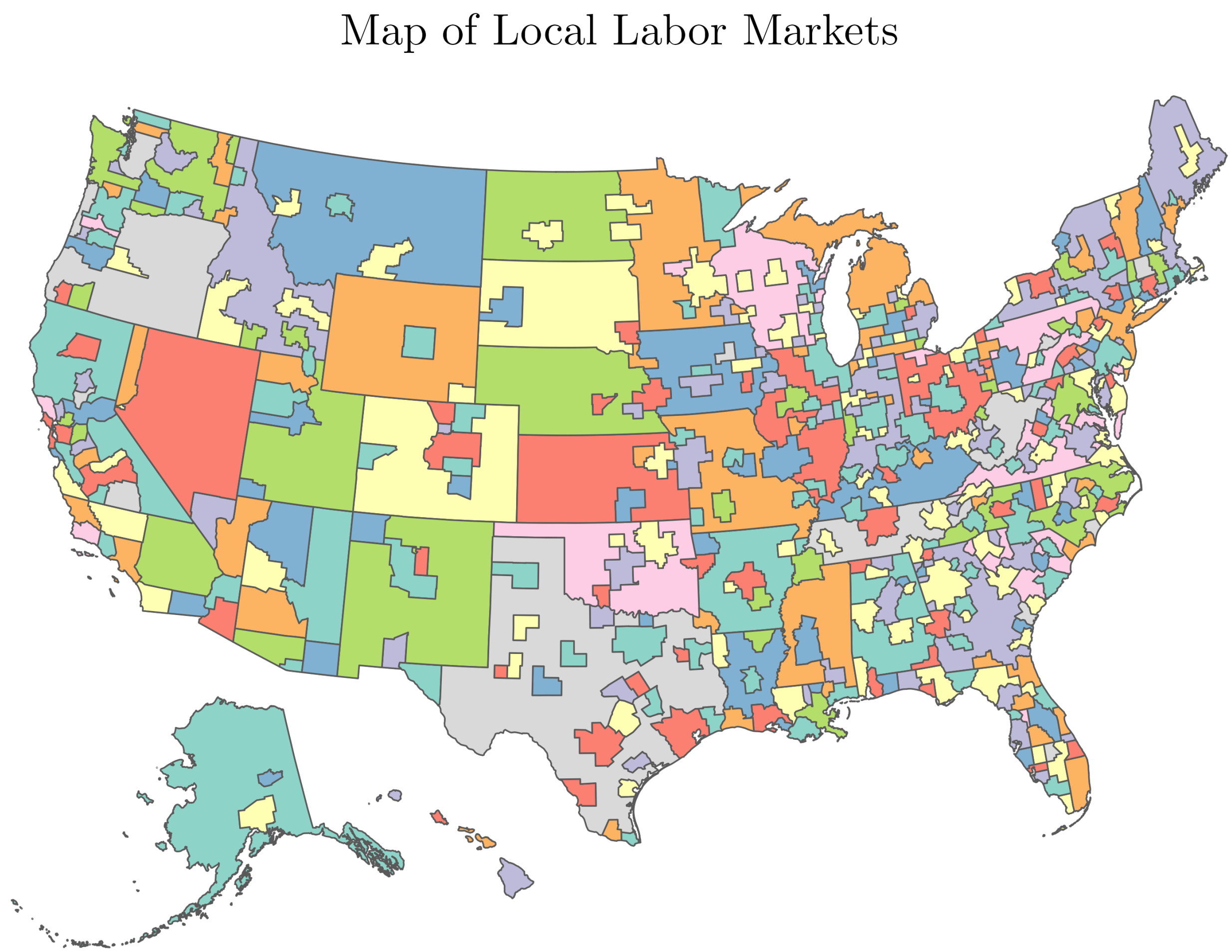

Figure 11: Local labor market definition used for analysis. LinkedIn aggregates user data to geographies that imperfectly approximate metropolitan statistical areas (Conzelmann et al. 2022). To maximize geographic granularity, I incorporate any metropolitan or micropolitan statistical areas included in LI. LinkedIn groups all non-metropolitan areas within a state together. Notice that non-metropolitan areas may not be contiguous, even though metropolitan labor markets are contiguous.

(c) Cleaning institutional string by removing numbers, punctuation, and standardizing capitalization.

(d) Identify obvious high school candidates, by searching for words like "hs", "senior", "high" in the institution's name. For candidate institutions like these, I automatically classify educational institutions as high schools.

(e) Group educational experiences by cleaned institution string, with the goal of discerning baccalaureate institutions from community colleges and high schools. I have the distribution of different degrees people attach to cleaned institution name string.

(f) Next I impute institution type based on the most common degree LI users attach to the string. Degree imputation hinges on the assumption that virtually every institution awarding high school diplomas in the United States does not award associate's, bachelor's, or graduate degrees. Thus, I estimate the non-missing values for the degree field using the most common degree in the institution described above.

(g) At this point, every educational experience a user lists should either have a user-supplied degree, or an imputed degree. Since I need both a high school and college for the analysis in the rest of the paper, I subset to users with both high school and college listed. I use the 2020-21 Education Demographic Geographic Estimates program, which provides school-level enrollment and geodata. I excluded any institutions with 0 enrollment in grade 12, namely elementary and middle schools. For baccalaurate institutions, I use the IPEDS 2021 Institutional Characteristics Complete Data File, which includes geodata, enrollment, completion, public status, and myriad other useful variables at the branch (not system) level.

(h) Nearly one-quarter of American high schools (10,988 out of 45,658) share a name with at least one other school. For example, the United States contains 89 Trinity Lutheran High Schools and 74 Central High Schools. In these cases, I assign users the closest high school to their college.

(i) If users list multiple baccalaureate institutions, I flag them as a transfer student. However, all subsequent analyses will rely on the last baccalaureate institution attended and their date of graduation. Every user will be attached to one high school and exactly one college, even if they attended multiple.

(j) I remove obvious cases of misrepresentation: users reporting a bachelor's degree or high school diploma from an institution that did not award them and completing a bachelor's degree in two years or less (including transfers across institutions)

3. Identifying post-graduation residence for jobs:

(a) Subset position files to users with both college and high school reported in educational file

(b) Identify earliest reported job on LinkedIn profile, which I use to impute college graduation date if the user does not supply one. I simply assume the user graduated in the year prior to starting the first job on their resume.

(c) The ultimate goal is a panel of residential locations by year. As a result, work experiences without start dates are useless, so I remove them. If the user failed to provide an end date, I use the date of the data pull: 17 July 2023. LinkedIn does not require start and end dates to be chronological, but I do. With the subset of work experiences that contain a start date and chronological end date, I can build a residential panel for each user.

(d) I also remove work experiences prior to their listed or imputed college graduation date. To standardize interpretation of results across different cohorts, I only care about migration in the years *following* college graduation.

(e) When users work list multiple work experiences within a year, I take whichever one they worked on January 1 or the earlier one.

(f) Finally I reshape the dataset, converting a wide dataset organized by start and end dates of work into a long data set of year $\times$ user residential locations. The locations will be strings, which I merge on a geographic crosswalk described in "Constructing crosswalks between LinkedIn geographies on labor markets"

6.2 Demographic Representation

Of course, LI suffers from limitations. Crucially, self-reporting of educational experiences leads to systematic under-reporting of certain experiences, namely non-completion of degrees (Kreisman, Smith, and Arifin 2021). Despite accounting for nearly half of all college attendees and roughly one-third of workers, non-completers often hide their post-secondary education. Individuals who attended for less than two years or attended non-selective institutions omit educational experiences at the highest rates (Kreisman, Smith, and Arifin 2021). To mitigate potential omission, I focus on college graduates, the educational group most likely to rely on LI (Dorn et al. 2025). Guillory and Hancock 2012 suggests the public nature of LI profiles discourages deceptive inflation of educational achievements and work history.

Dorn et al. 2025 relies on the same Revelio database. Figure A2 in Dorn et al. 2025 characterizes the sample relative to the 23-64 labor force, according to the 2019 ACS. Note that Dorn et al. 2025 restrict the sample by age rather than college graduation cohort, without imposing the costly requirement of a high school and college in the same profile. The selection among LI users relative to ACS seems most pronounced with respect to undergraduate major and occupational grouping, with the sample largely balanced across regions (Dorn et al. 2025). Panel C of Figure A2 in shows that workers with business and STEM degrees are over-represented among LI users by 7 and 5 p.p. respectively relative to the American prime-age labor force in 2019. The differences across majors probably reflect the distinct occupational distributions of different majors. For example, business majors probably enter managerial and financial occupations at higher rates than nursing majors, who tend to work in health care occupations. Thus, the over-representation of business and STEM majors in Panel C tracks with Panel B, which characterizes occupational composition. Panel B suggests workers in managerial, technical, and creative occupations use LI more frequently than health care, education, and construction workers. The differences in usage might reflect differences in job search frequency. LI is primarily a job search tool, so it offers the greatest utility for workers frequently switching jobs. Workers lean on their networks when searching for jobs (Krolikowski, Zabek, and Coate 2020), so workers with frequent job switches will have broader and larger networks. Yet workers in health care and education tend to stay at the same establishment longer than many of their fellow workers. Furthermore, non-tradable sectors like education and health care are present in large numbers in every local labor market of the United States, limiting the importance of social connections.

While no dataset is perfect, experimental evidence from Guillory and Hancock 2012 and validation exercises in Conzelmann et al. 2023; Berry, Maloney, and Neumark 2024; Dorn et al. 2025 demonstrate LI's value as a tool to the study the labor market.

6.2.1 Baccalaureate Institutional Grouping

I use the following system for characterizing baccalaureate institutions:

Flagship Public institutions that are federally-recognized land-grant college or university, members of the American Association of Universities, or both[28]

Private, Non-Profit Institutions classified as private and not-for-profit in IPEDS.

Private, For-Profit Institutions classified as private and for-profit in IPEDS.

RPU Public institutions not satisfying the requirements for Flagship

28. The Department of Education does not officially classify institutions as "flagship". Nonetheless, I wanted to distinguish the large research universities that often play an outsize role in policy debates about higher education (Bound et al. 2019). Bound et al. 2019 defines flagship using AAU members alone. I found the definition overly restrictive since the AAU only counts eight members across the South and six in the west outside of California. I wanted at least one "flagship" institution in every state. By taking the union of Land Grant and AAU members, I incorporate the geographic spread of the Land Grant institutions, which appear in all 50 states and DC, without overlooking prominent public institutions established prior to the Morrill Act, such as the University of Virginia and University of Michigan. Table 19 in the Appendix lists all institutions classified as Flagship.

The governance structures of public universities allow state policymakers to effect change through appropriations and personnel decisions. Private non-profit universities face similar pressures, though state government wields fewer policy levers for enacting reform.[29] Much of the analysis focuses on public universities, delineating between **public flagship** (PFs) institutions and **regional public universities** (RPUs) with greater emphases on undergraduate education. Table 6 characterizes the institutional groupings in greater detail.

Table 6: Summary statistics by institutional grouping

Measure	Public Flagship	Private Non-Profit	Regional Public
Agriculture	4.5	0.8	1.3
Business	14.7	21.4	19.9
Education	7.8	7.2	9.6
Engineering	15.6	8.0	8.8
Humanities	11.0	15.2	13.5
Physical Sciences	11.9	8.4	7.1
Services	2.7	4.1	6.5
Social Sciences	23.6	20.5	20.5
Healthcare	5.2	10.2	9.8
Other	2.9	4.2	3.2
Acceptance Rate	63.5	61.6	70.5
In-State	75.7	52.4	87.7
Mobility Rate	2.0	1.9	2.0
Parent Mean Income (2013)	146,000	191,100	107,900
Kid Mean Income (2013)	61,300	64,200	48,000
Parents From Bottom 60%	32.5	32.6	41.6
Parents From Top 20%	23.5	22.7	26.5
Parents From Top 1%	2.5	4.9	1.1
Kids in Bottom 60%	34.0	35.8	41.1
Kids in Top 20%	23.1	22.3	27.4
Kids in Top 1%	3.3	4.2	1.4
State Appropriations % of Revenue	17.2	0.3	23.9
Tuition % of Revenue	18.0	33.1	23.4
Completed in 4 Years	44.7	51.8	25.1
Completed in 6 Years	69.7	64.8	49.8
Completed in 8 Years	72.1	66.3	53.2

Data on majors, acceptance rates, finances, and completion rates sourced from IPEPS vintages 2000-2022. Mobility and income data from Chetty et al. 2017. PFs differ from RPUs in substantive ways. PFs award more engineering and fewer healthcare or business degrees. The parental income distribution at PFs resembles the distribution at Private colleges more closely than RPUs, with more students from the top two deciles (Chetty et al. 2017). Likewise, PF outcomes mirror Private institutions, rather than RPUs. Graduates of PFs earn more, with an income distribution nearly identical for PFs and Privates (Chetty et al. 2017). PFs boast degree completion rates 20 percentage points higher than their public peers, even 8 years after matriculation. Finally, the institutions receive revenue from different sources. PFs rely less on state appropriations and tuition than RPUs.

LinkedIn users are not a random sample of college graduates (Dorn et al. 2025; Berry, Maloney, and

29. The incentives of private for-profit universities differ from both public and private non-profit institution types. For-profit institutions' reliance on virtual educational services complicates measurement of brain drain, since degree seekers can live anywhere in the US.As a result, I exclude graduates of for-profit colleges from the analysis.

Neumark 2024; Conzelmann et al. 2022). Figure 12 underscores the lack of representativeness. In general, colleges awarding more degrees tend to appear more frequently in the data. However, RPUs are under-represented in the data, clustered tightly to the left of the dashed $y = x$ line. The under-representation of RPU graduates and over-representation of Private college graduates reflects a variety of factors, including major and occupation choice.

6.2.2 Occupation

The occupational composition of the LI sample in Dorn et al. 2025 differs substantially from ACS benchmarks. LI over-represents workers in creative, financial, technical, or managerial occupations, with the disparity most pronounced in managerial and business occupations. Workers in eduction, health care, construction, and non-managerial office positions are under-represented among LI users. Relative to Dorn et al. 2025, I use a smaller sample concentrated in a narrower set of birth cohorts. To explore occupation in greater detail, I use the six-digit SOCs attached to each job listed on LI, taking the most frequent SOC for each individual. Using occupational prestige data from Condon and Hughes 2022, I crosswalk 6-digit SOCs to one of three evenly-sized bins. Each bin represents one-third of the college-educated labor force, with education $\times$ occupation cells estimated with 2022 BLS data.[30] Two digit SOCs are quite broad, often encompassing dozens of six-digit occupations. As a result, the broad two-digit categories often encompass occupations with different levels of prestige. Figure 13 aggregates upward to combinations of occupational bins and two-digit SOCs. The LI analytical sample holds up well; increasing the occupational group's share of the analytical sample by 1 percent increases the BLS share by 0.882 percent. The plot shows values of occupational groups on both sides of the $y = x$ line, suggesting systematic over-representation of broad occupational categories is not biased with respect to the prestige bins.

Figure 14 characterizes the joint distribution of college graduates by institutional grouping and occupational prestige. RPU graduates are more likely to work in low-prestige occupations. Compared to PF alumni, RPU graduates are 5 p.p. more likely to work in low-prestige occupations, with the gap persisting as workers gain experience. The gap grows to 10 p.p. relative to Private colleges. Table 12 features the largest six-digit SOCs in each bin; the five largest occupations account for 6.4 percent of college-educated workers but only 1.2 percent of the sample. The x-axis on Figure 14 plots bins of career experience, based on years since college graduation. Within an institutional group, shares of workers in each bin remain stable even as workers gain experience. The career progression observed in LI data largely occurs within the same bin or movements between bins are balanced. Regardless, occupational prestige is not just proxying for age,

30. Employment by detailed occupation linked with Educational attainment by detailed occupation enables estimation of the number of college graduates in each occupation. Many college graduates work in occupations that do not require a bachelor's degree. For example, retail salespeople (SOC: 41-2031) outnumbered lawyers (SOC:23-1011) in 2022.

but capturing a distinct axis of variation within the data.

Graduates of RPUs are under-represented and Private college alumni over-represented in the analytical sample. LinkedIn coverage varies substantially at the six-digit SOC, though grouping to two-digit SOCs and occupational bins attenuates the variation. Grouping by occupational bins also provides stabilizes cross-sectional comparisons across institutional groupings. Graduates of RPUs are more likely to work in low-prestige jobs, while PFs and Private colleges exhibit similar distributions of occupational prestige.

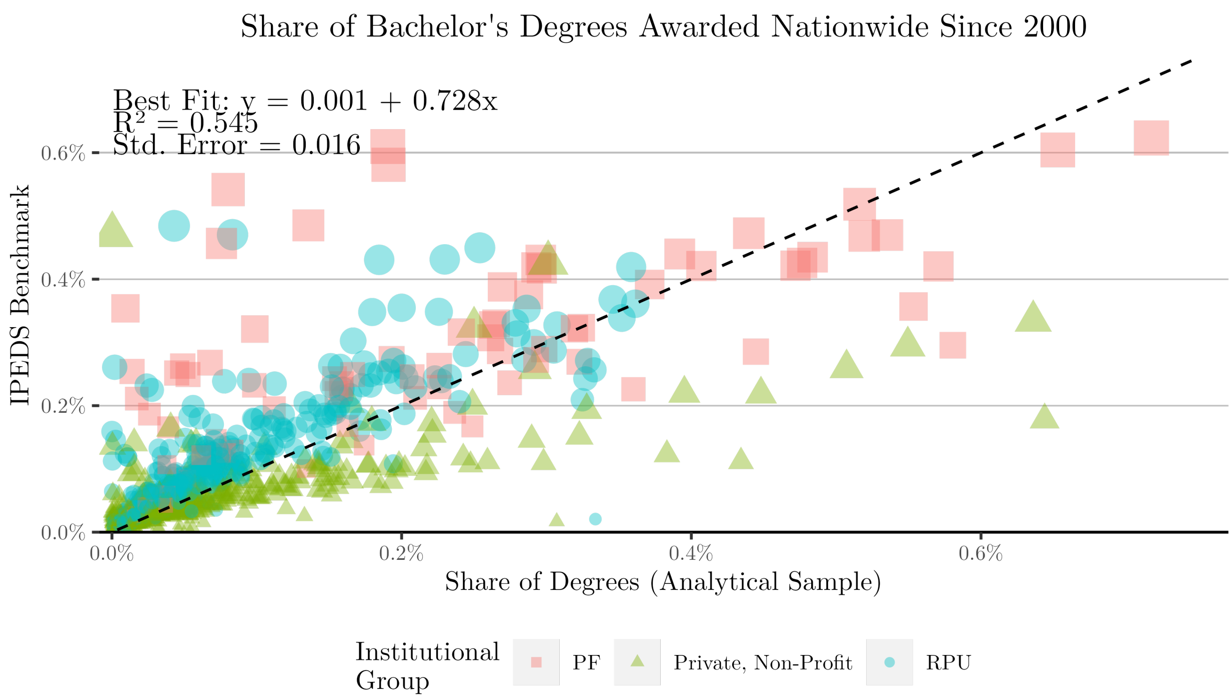

Figure 12: The analytical sample over-represents alumni of Private colleges and under-represents RPUs. The x-axis reports an institution's share of total degrees in the sample, while the y-axis reports the institution's share of total degrees awarded since 2000. If the data were perfectly representative, every point would fall on the black dotted line. Despite potential issues with under-representation of RPU graduates, colleges that awarded more degrees tend appear more frequently in the data.

Table 6: The most common 6-digit SOC codes by prestige bins

High Prestige				
Occupation Title	SOC	Graduates	LI Share	BLS Share
Registered nurses	29-1141	2237.5	0.56	3.74
Software developers	15-1252	1462.0	1.66	2.44
Lawyers	23-1011	841.0	1.62	1.41
Business operations specialists	13-1199	747.9	0.18	1.25
Managers, all other	11-9199	741.3	0.78	1.24
Medium Prestige				
Occupation Title	SOC	Graduates	LI Share	BLS Share
General and operations managers	11-1021	1662.6	0.41	2.78
Accountants and auditors	13-2011	1376.1	0.70	2.30
Elementary school teachers	25-2021	1374.7	0.67	2.30
Secondary school teachers	25-2031	1035.0	0.51	1.73
Management analysts	13-1111	813.6	0.74	1.36
Low Prestige				
Occupation Title	SOC	Graduates	LI Share	BLS Share
Retail salespersons	41-2031	1021.1	0.70	1.71
Customer service representatives	43-4051	827.3	1.14	1.38
Fast food and counter workers	35-3023	687.1	0.58	1.15
Sales representatives	41-4012	669.9	0.13	1.12
Office clerks, general	43-9061	611.2	0.00	1.02

Many college graduates work in occupations with lower prestige and occupational status than popular perception suggests. For example, the number of secondary school teachers approximates the number of retail salespeople with a bachelor's degree. The low-prestige generally bin consists of frontline service workers and production jobs. By contrast, the medium and high prestige bins include more occupations traditionally associated with college graduates, often involving more abstract and analytical work.

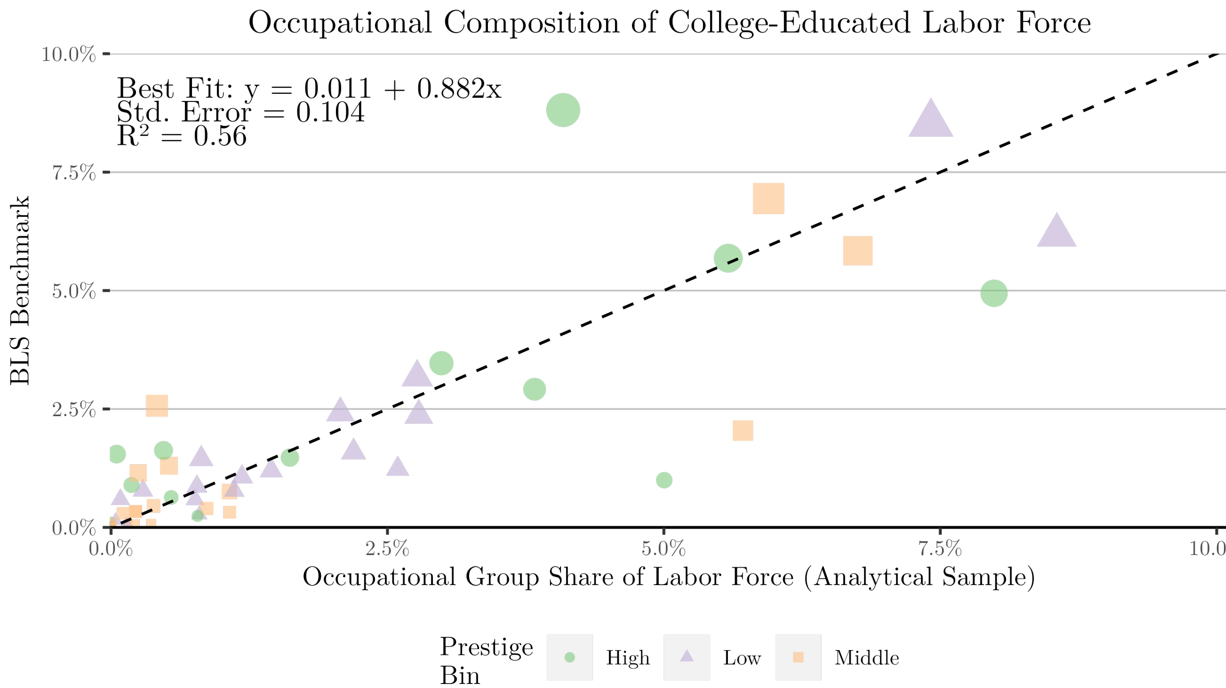

Figure 13: Each point on the plot represents 2-digit SOC × prestige bin cell. The x-axis reports the occupational group-prestige cell's share of observations in the analytical sample, which the y-axis shows the cell's share of the college-educated labor market according to 2022 BLS data. For example, the purple triangle in the upper right above the dotted light represents low-prestige Sales jobs (SOC: 41), while the triangle below it plots Office and Administrative Support (SOC: 43). The sample's representativeness varies greatly across cells, but less across prestige bins that group cells together.

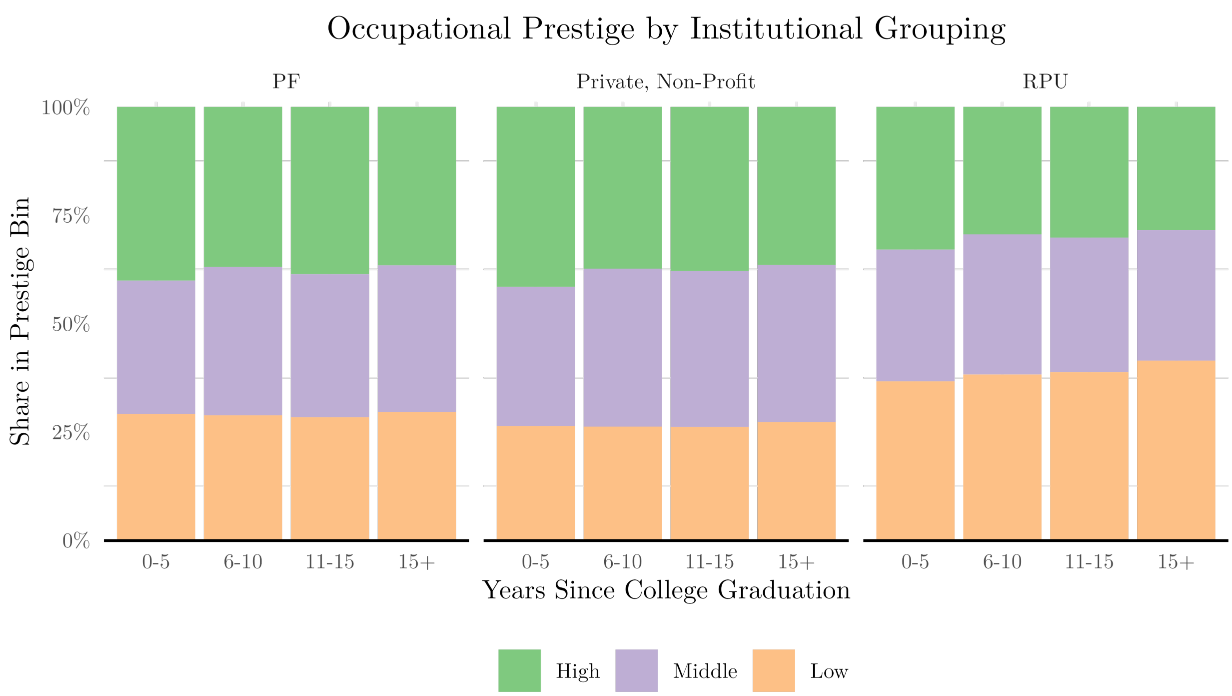

Figure 14: Observations were assigned a prestige bin based on the longest-held occupation listed in their LI profile. Conditional on institutional grouping, the shares of workers in each prestige bin remains constant regardless of the time elapsed since college graduation. Life-cycle models suggest workers with less experience will cluster in lower-prestige occupations. Even as workers progress in their careers, the shares of workers in each prestige bin barely moves. Instead, the differences observed across baccalaureate institutions drives the variation. On average, graduates of regional public universities work in lower prestige jobs, reflecting the lower earnings for graduates seen in Table 1.

6.2.3 Age

LinkedIn users are younger than the working population, so analytical sample over-represents workers born between 1977 and 1992. Figure 15 reports the age distribution of both the sample and college graduates in the labor force according to the 2022 American Community Survey. Much of the sample falls into the youngest quartile of the labor force (those born after 1987). Accordingly, Figure 16 shows the sample includes more than 3 percent of all college-educated workers born in the late 1980s and 1990s, but fewer than 2 percent for cohorts born before 1970.

If someone included a graduation date from high school or college, I subtracted 18 to estimate birth the year. Compared to other graduation dates, high school graduations are the best predictor of age because most graduates are 17 or 18. By contrast, more than 27 percent of first-time full-time students who entered college in 2014 graduated in more than four years, accounting for roughly one-third of all graduates from the cohort. Clearly, college graduation dates vary more than high school graduations for a given age cohort. I circumvent the problem by taking the start year of a bachelor's degree included on LinkedIn. For first-time full-time students, start dates should not exhibit such wide variation. If a user includes a bachelor's degree with a start year, I subtract 18.

If individuals attended multiple institutions in the United States for bachelor's degrees (i.e. transfer students), I estimate age using the start date of the earliest institution and measure post-graduation behavior based on the end date at the latest institution.

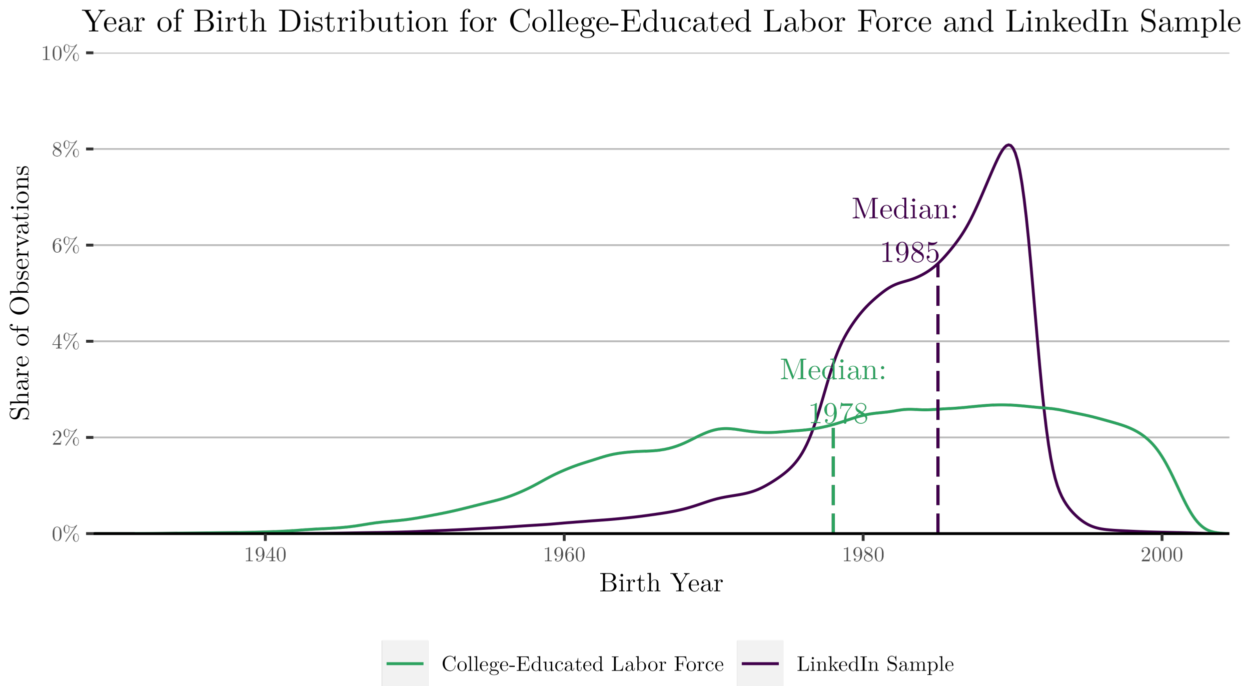

Figure 15: Age distribution of analytical sample compared to college-educated workforce according to 2022 American Community Survey microdata. The sample over-represents users born between 1977 and 1992. Likewise, the median observation in the analytical sample is 7 years younger than the median college-educated worker (1978 vs. 1985). The older half of the college-educated workforce represents less than a quarter of the sample.

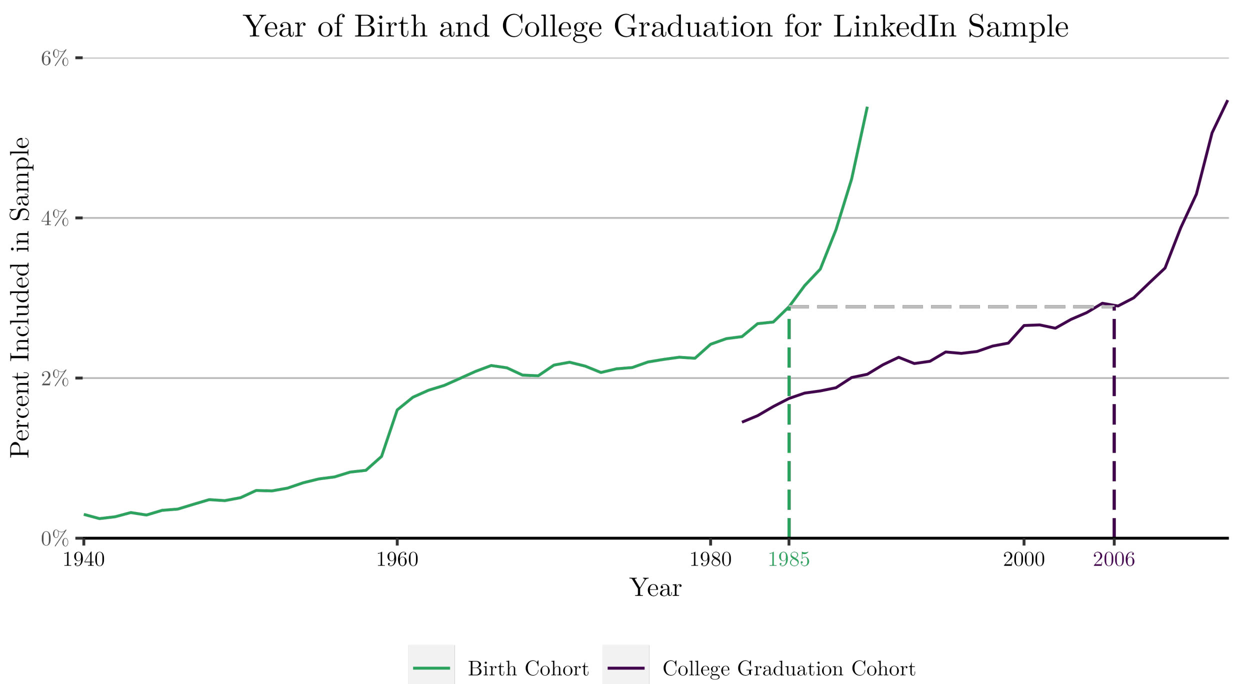

Figure 16: The y-axis measures the LI user sample as a share of all college graduates, by birth and college graduation year. Estimates of birth years come from 2022 American Community Survey microdata and college graduation dates from the 2022 Digest of Education Statistics. Clearly, the analytical sample primarily includes younger workers. Cohorts born in the 1980s and 1990s appear at far higher rates than older workers. Coverage improves dramatically for cohorts born after 1985 and graduating college during or after the Great Recession. The figure also constitutes a robustness check on the process for estimating birth cohorts outlined in Section 6.2.3. Regardless of year of birth, users in the sample graduate an average of 23 years later, reflecting the fact that bachelor's completion takes more than four years for many.

7 Additional Results

7.1 Complier Means Table

Table 7: Characteristics of compliers

Characteristic	Compliers	Std. Error	Sample	All Migrants
Regional Public University	0.39	0.02	0.36	0.36
Public Flagship	0.23	0.01	0.22	0.23
Attended In-State College	0.64	0.03	0.49	0.51
Transferred Between Colleges	0.21	0.00	0.25	0.24
Completed Degree by 24	0.80	0.00	0.73	0.75
Low-Prestige Occp.	0.32	0.01	0.29	0.31
Medium-Prestige Occp.	0.34	0.00	0.34	0.34
RPU, In-State	0.32	0.02	0.24	0.25
PF, In-State	0.15	0.01	0.12	0.13
Private, In-State	0.18	0.01	0.14	0.14
Return Within 5 Years	0.77	0.00	0.23	0.30
Return in 5-10 Years	0.22	0.00	0.06	0.08
Bottom LD Quartile	0.24	0.02	0.27	0.26
Middle LD Quartiles	0.49	0.04	0.47	0.48
Top LD Quartile	0.27	0.03	0.26	0.26
High-Amenity LM	0.77	0.04	0.71	0.74
Low-Amenity LM	0.23	0.04	0.29	0.26

Compliers are college graduates induced to return to their high school labor market by labor demand shocks. Estimation of complier means relies on the Kappa theorem in Abadie 2003. The leftmost column reports shares of compliers across a variety of characteristics, with the next column reporting White standard errors for the complier shares. The Sample column includes all college graduates who moved at some point after graduating college between 2000 and 2013. The All Migrants column includes anybody graduating between 2000 and 2013. Compliers tend to graduate from Regional Public Universities and baccalaureate institutions in the same state as their high schools. In-state graduates are nearly two-thirds of compliers, despite accounting for half of the sample. College graduates who return to origin labor markets mirror the occupational composition of the labor market, with each occupational prestige bin accounting for about one-third of compliers. Figure 17 visualizes the data.

7.2 First-Stage Regression

7.3 Alternative Specifications of Second Stage

7.3.1 Migration Within Five Years

Table 8: First-stage of least-squares estimation

Dependent Variables:	HS LD Shock		College LD Shock
Model:	(1)	(2)	(3)
Variables			
HS LD Shock	0.9767***	0.9764***	0.0023***
	(0.0038)	(0.0057)	(0.0008)
College LD Shock		0.0034***	0.9842***
		(0.0007)	(0.0023)
Fit statistics			
Observations	77,695	396,546	396,546
R^2	0.96905	0.97046	0.97641
F-test (1st stage)	1,591,231.0	4,795,502.1	6,441,059.9

Clustered (hs_cbsa_code) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

White standard errors, clustered at the labor market level, reported in parentheses. Each column estimates estimates Equation 6 (equivalent to Model (5) in any of the second-stage regressions). The model controls for institutional grouping, residency, major, transfer, on-time completion controls, and college graduation cohort fixed effects. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is the endogenous variable: annual growth in the labor market. Section 4.1 explains the first-stage estimation process in greater detail.

Table 9: Probability of returning to labor market containing college within 5 years of college graduation

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	0.026	−0.011	−0.007	−0.004	−0.056*	−0.053	−0.012	−0.006	−0.002	−0.066*
	(0.032)	(0.032)	(0.032)	(0.032)	(0.028)	(0.034)	(0.035)	(0.035)	(0.034)	(0.029)
College LD Shock	0.424***	0.402***	0.410***	0.390***	0.359***	0.438***	0.483***	0.489***	0.464***	0.428***
	(0.025)	(0.029)	(0.029)	(0.028)	(0.027)	(0.026)	(0.031)	(0.032)	(0.030)	(0.029)
Private			0.081***	0.080***	0.085***			0.081***	0.080***	0.084***
			(0.006)	(0.006)	(0.005)			(0.006)	(0.006)	(0.005)
RPU			0.129***	0.127***	0.130***			0.129***	0.127***	0.130***
			(0.006)	(0.006)	(0.006)			(0.006)	(0.006)	(0.006)
Private × In-State			−0.017	−0.028*	−0.023			−0.017	−0.028*	−0.023
			(0.014)	(0.014)	(0.013)			(0.014)	(0.014)	(0.013)
RPU × In-State			−0.077***	−0.079***	−0.076***			−0.077***	−0.079***	−0.076***
			(0.014)	(0.014)	(0.013)			(0.014)	(0.014)	(0.013)
College In State			0.056***	0.070***	0.064***			0.056***	0.070***	0.064***
			(0.012)	(0.012)	(0.012)			(0.012)	(0.012)	(0.012)
BA by Age 24				−0.072***	−0.070***				−0.071***	−0.070***
				(0.005)	(0.004)				(0.005)	(0.004)
Transfer Student				0.023***	0.023***				0.022***	0.023***
				(0.004)	(0.003)				(0.004)	(0.003)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34
Num. obs.	365023	365023	365023	365023	365023	365023	365023	365023	365023	365023

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

First-stage partial F-statistic for high school LD shock is 4795502 ($R^2 = 0.055$) and 6441060 for college shock ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor markets as endogenous independent variables, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) incorporates college graduation cohort fixed effects, Model (3) controls for institutional grouping and residency, and Model (4) adds major, transfer, and on-time completion controls. Model (5) estimates Equation 6. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate ever returned to the labor market containing their **college** between the first and fifth years after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to high school labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. Table 8 reports results of the first-stage estimation process described in Section 4.1.

Table 10: Probability of returning to labor market containing high school within 5 years of college graduation

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	0.332***	0.181***	0.180***	0.176***	0.169***	0.215***	0.210***	0.212***	0.207***	0.185***
	(0.062)	(0.036)	(0.034)	(0.034)	(0.034)	(0.050)	(0.042)	(0.042)	(0.041)	(0.036)
College LD Shock	0.005	−0.112***	−0.112***	−0.099***	−0.094***	−0.139***	−0.129***	−0.123***	−0.107***	−0.104***
	(0.028)	(0.031)	(0.027)	(0.027)	(0.018)	(0.029)	(0.032)	(0.029)	(0.028)	(0.019)
Private			−0.001	−0.002	−0.003			−0.000	−0.002	−0.003
			(0.007)	(0.007)	(0.004)			(0.007)	(0.007)	(0.004)
RPU			−0.027***	−0.025***	−0.020***			−0.027***	−0.025***	−0.020***
			(0.007)	(0.007)	(0.005)			(0.007)	(0.007)	(0.005)
Private × In-State			−0.047***	−0.040***	−0.050***			−0.047***	−0.040***	−0.050***
			(0.009)	(0.009)	(0.008)			(0.009)	(0.009)	(0.008)
RPU × In-State			0.021	0.019	0.011			0.021	0.019	0.011
			(0.011)	(0.011)	(0.007)			(0.011)	(0.011)	(0.007)
College In State			0.086***	0.075***	0.087***			0.086***	0.075***	0.087***
			(0.012)	(0.013)	(0.008)			(0.012)	(0.012)	(0.008)
BA by Age 24				0.065***	0.062***				0.065***	0.062***
				(0.007)	(0.006)				(0.007)	(0.006)
Transfer Student				−0.000	−0.001				−0.000	−0.001
				(0.002)	(0.002)				(0.002)	(0.002)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13
Num. obs.	365023	365023	365023	365023	365023	365023	365023	365023	365023	365023

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

First-stage partial F-statistic for high school LD shock is 4795502 ($R^2 = 0.055$) and 6441060 for college shock ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor markets as endogenous independent variables, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) incorporates college graduation cohort fixed effects, Model (3) controls for institutional grouping and residency, and Model (4) adds major, transfer, and on-time completion controls. Model (5) estimates Equation 6.White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate ever returned to the labor market containing their **high school** between the first and fifth years after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to high school labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. Table 8 reports results of the first-stage estimation process described in Section 4.1.

Table 11: Probability of returning to labor market containing high school and college within 5 years of college graduation

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	0.719***	0.081	0.085	0.080	0.062	0.170	0.126	0.131	0.126	0.076
	(0.172)	(0.087)	(0.086)	(0.088)	(0.080)	(0.102)	(0.098)	(0.097)	(0.098)	(0.085)
Private			−0.027	−0.026	−0.040			−0.027	−0.026	−0.040
			(0.053)	(0.051)	(0.041)			(0.053)	(0.051)	(0.041)
RPU			−0.010	−0.004	0.015			−0.009	−0.004	0.015
			(0.057)	(0.057)	(0.050)			(0.057)	(0.057)	(0.050)
Private, Non-Profit × In-State			0.040	0.038	0.045			0.040	0.038	0.045
			(0.055)	(0.053)	(0.041)			(0.055)	(0.053)	(0.041)
RPU × In-State			0.022	0.020	0.003			0.022	0.020	0.003
			(0.058)	(0.057)	(0.049)			(0.058)	(0.057)	(0.049)
College In State			0.002	0.006	0.014			0.002	0.007	0.014
			(0.057)	(0.056)	(0.044)			(0.057)	(0.056)	(0.044)
BA by Age 24				0.035***	0.035***				0.035***	0.035***
				(0.006)	(0.006)				(0.006)	(0.006)
Transfer Student				0.027***	0.023***				0.026***	0.023***
				(0.006)	(0.006)				(0.006)	(0.006)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38
Num. obs.	65245	65245	65245	65245	65245	65245	65245	65245	65245	65245

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

First-stage partial F-statistic for high school LD shock is 4795502 ($R^2 = 0.055$) and 6441060 for college shock ($R^2 = 0.055$). First-stage F-statistic for high school and college LD shock is 1591231 ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor market as endogenous independent variables, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) incorporates college graduation cohort fixed effects, Model (3) controls for institutional grouping and residency, and Model (4) adds major, transfer, and on-time completion controls. Model (5) estimates Equation 6. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate ever returned to the labor market containing their **high school and college** between the first and fifth years after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to high school labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. Table 8 reports results of the first-stage estimation process described in Section 4.1.

7.3.2 Residence in Years 1 and 10 Only as Dependent Variable

Table 12: Probability of returning to labor market containing college by 10th year since college graduation

	OLS					**IV**				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	−0.076**	−0.002	0.000	0.007	−0.030	−0.067*	−0.004	−0.001	0.008	−0.039
	(0.026)	(0.027)	(0.027)	(0.026)	(0.022)	(0.028)	(0.029)	(0.029)	(0.028)	(0.024)
College LD Shock	0.261***	0.322***	0.329***	0.307***	0.283***	0.339***	0.390***	0.394***	0.368***	0.339***
	(0.018)	(0.022)	(0.022)	(0.021)	(0.021)	(0.021)	(0.024)	(0.025)	(0.023)	(0.023)
Private			0.075***	0.074***	0.076***			0.075***	0.073***	0.075***
			(0.006)	(0.006)	(0.005)			(0.006)	(0.006)	(0.005)
RPU			0.122***	0.118***	0.120***			0.122***	0.118***	0.120***
			(0.006)	(0.006)	(0.005)			(0.006)	(0.006)	(0.005)
Private × In-State			0.002	−0.011	−0.008			0.002	−0.011	−0.007
			(0.013)	(0.013)	(0.012)			(0.013)	(0.013)	(0.012)
RPU × In-State			−0.069***	−0.070***	−0.069***			−0.069***	−0.070***	−0.069***
			(0.013)	(0.013)	(0.012)			(0.013)	(0.013)	(0.012)
College In State			0.042***	0.059***	0.056***			0.042***	0.059***	0.056***
			(0.011)	(0.011)	(0.011)			(0.011)	(0.011)	(0.011)
BA by Age 24				−0.094***	−0.093***				−0.093***	−0.093***
				(0.004)	(0.003)				(0.004)	(0.003)
Transfer Student				0.008*	0.008*				0.008*	0.008*
				(0.003)	(0.003)				(0.003)	(0.003)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23
Num. obs.	333422	333422	333422	333422	333422	333422	333422	333422	333422	333422

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

First-stage partial F-statistic for high school LD shock is 4795502 ($R^2 = 0.055$) and 6441060 for college shock ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor markets as endogenous independent variables, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) incorporates college graduation cohort fixed effects, Model (3) controls for institutional grouping and residency, and Model (4) adds major, transfer, and on-time completion controls. Model (5) estimates Equation 6. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate resided to the labor market containing their **college** in the tenth year after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to high school labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. Table 8 reports results of the first-stage estimation process described in Section 4.1.

Table 13: Probability of returning to labor market containing high school by 10th year since college graduation

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	0.214***	0.143***	0.143***	0.141***	0.122***	0.137***	0.167***	0.169***	0.166***	0.135***
	(0.039)	(0.029)	(0.028)	(0.028)	(0.025)	(0.030)	(0.034)	(0.034)	(0.033)	(0.027)
College LD Shock	0.021	−0.030	−0.029	−0.021	−0.024	−0.072**	−0.041	−0.035	−0.025	−0.030*
	(0.021)	(0.025)	(0.022)	(0.022)	(0.015)	(0.022)	(0.025)	(0.022)	(0.022)	(0.015)
Private, Non-Profit			0.002	0.001	0.002			0.002	0.001	0.002
			(0.006)	(0.006)	(0.003)			(0.006)	(0.006)	(0.003)
RPU			−0.019***	−0.018***	−0.013***			−0.019***	−0.018***	−0.013***
			(0.005)	(0.005)	(0.004)			(0.005)	(0.005)	(0.004)
Private, Non-Profit *times* In-State			−0.038***	−0.034***	−0.042***			−0.038***	−0.034***	−0.042***
			(0.008)	(0.008)	(0.006)			(0.008)	(0.008)	(0.006)
RPU × In-State			0.016	0.014	0.007			0.016	0.014	0.007
			(0.009)	(0.009)	(0.006)			(0.009)	(0.009)	(0.006)
College In State			0.061***	0.054***	0.062***			0.061***	0.054***	0.062***
			(0.009)	(0.009)	(0.005)			(0.009)	(0.009)	(0.005)
BA by Age 24				0.041***	0.040***				0.041***	0.040***
				(0.005)	(0.004)				(0.005)	(0.004)
Transfer Student				0.001	0.001				0.001	0.001
				(0.002)	(0.002)				(0.002)	(0.002)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Num. obs.	333422	333422	333422	333422	333422	333422	333422	333422	333422	333422

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

First-stage partial F-statistic for high school LD shock is 4795502 ($R^2 = 0.055$) and 6441060 for college shock ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor markets as endogenous independent variables, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) incorporates college graduation cohort fixed effects, Model (3) controls for institutional grouping and residency, and Model (4) adds major, transfer, and on-time completion controls. Model (5) estimates Equation 6. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate resided to the labor market containing their **high school** in the tenth year after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to high school labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. Table 8 reports results of the first-stage estimation process described in Section 4.1.

Table 14: Probability of returning to labor market containing high school and college by 10th year since college graduation

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	0.220***	0.086	0.091	0.086	0.079	0.060	0.117*	0.120*	0.115*	0.091
	(0.051)	(0.056)	(0.054)	(0.055)	(0.049)	(0.044)	(0.058)	(0.057)	(0.058)	(0.050)
Private			−0.040	−0.038	−0.035			−0.040	−0.038	−0.035
			(0.033)	(0.032)	(0.024)			(0.033)	(0.032)	(0.024)
RPU			−0.013	−0.008	0.005			−0.013	−0.007	0.005
			(0.028)	(0.027)	(0.024)			(0.028)	(0.027)	(0.024)
Private × In-State			0.062	0.058	0.052*			0.062	0.058	0.052*
			(0.035)	(0.033)	(0.025)			(0.035)	(0.033)	(0.025)
RPU × In-State			0.021	0.016	0.006			0.021	0.016	0.006
			(0.031)	(0.030)	(0.026)			(0.031)	(0.030)	(0.026)
College In State			−0.043	−0.038	−0.027			−0.042	−0.037	−0.027
			(0.030)	(0.028)	(0.026)			(0.030)	(0.028)	(0.026)
BA by Age 24				0.005	0.005				0.005	0.005
				(0.005)	(0.005)				(0.005)	(0.005)
Transfer Student				0.010*	0.008				0.010*	0.008
				(0.005)	(0.005)				(0.005)	(0.005)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18
Num. obs.	66149	66149	66149	66149	66149	66149	66149	66149	66149	66149

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

First-stage F-statistic for high school and college LD shock is 1591231 ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor market as endogenous independent variables, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) incorporates college graduation cohort fixed effects, Model (3) controls for institutional grouping and residency, and Model (4) adds major, transfer, and on-time completion controls. Model (5) estimates Equation 6. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate resided to the labor market containing their **high school and college** in the tenth year after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to high school labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. Table 8 reports results of the first-stage estimation process described in Section 4.1.

7.3.3 Residence in Years 1 and 5 Only as Dependent Variable

Table 15: Probability of returning to labor market containing college by 5th year since college graduation

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	−0.081**	−0.003	0.000	0.005	−0.039	−0.095**	0.001	0.005	0.013	−0.045
	(0.027)	(0.029)	(0.030)	(0.029)	(0.025)	(0.031)	(0.031)	(0.032)	(0.031)	(0.026)
College LD Shock	0.314***	0.381***	0.390***	0.367***	0.340***	0.376***	0.456***	0.463***	0.435***	0.402***
	(0.020)	(0.025)	(0.025)	(0.024)	(0.024)	(0.022)	(0.027)	(0.028)	(0.027)	(0.026)
Private, Non-Profit			0.083***	0.082***	0.084***			0.082***	0.081***	0.083***
			(0.006)	(0.006)	(0.005)			(0.006)	(0.006)	(0.005)
RPU			0.129***	0.125***	0.127***			0.129***	0.125***	0.127***
			(0.006)	(0.006)	(0.005)			(0.006)	(0.006)	(0.005)
Private × In-State			−0.006	−0.020	−0.018			−0.006	−0.020	−0.018
			(0.014)	(0.014)	(0.013)			(0.014)	(0.014)	(0.013)
RPU × In-State			−0.071***	−0.073***	−0.073***			−0.071***	−0.073***	−0.073***
			(0.014)	(0.013)	(0.013)			(0.013)	(0.013)	(0.013)
College In State			0.050***	0.067***	0.066***			0.050***	0.067***	0.066***
			(0.011)	(0.011)	(0.011)			(0.011)	(0.011)	(0.011)
BA by Age 24				−0.094***	−0.094***				−0.094***	−0.093***
				(0.004)	(0.004)				(0.004)	(0.004)
Transfer Student				0.011***	0.012***				0.011***	0.011***
				(0.003)	(0.003)				(0.003)	(0.003)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Num. obs.	325544	325544	325544	325544	325544	325544	325544	325544	325544	325544

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

First-stage partial F-statistic for high school LD shock is 4795502 ($R^2 = 0.055$) and 6441060 for college shock ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor markets as endogenous independent variables, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) incorporates college graduation cohort fixed effects, Model (3) controls for institutional grouping and residency, and Model (4) adds major, transfer, and on-time completion controls. Model (5) estimates Equation 6. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate resided to the labor market containing their **college** in the fifth year after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to high school labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. Table 8 reports results of the first-stage estimation process described in Section 4.1.

Table 16: Probability of returning to labor market containing high school by 5th year since college graduation

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	0.245^{***}	0.159^{***}	0.159^{***}	0.156^{***}	0.141^{***}	0.174^{***}	0.178^{***}	0.179^{***}	0.176^{***}	0.149^{***}
	(0.043)	(0.028)	(0.027)	(0.027)	(0.025)	(0.037)	(0.031)	(0.031)	(0.031)	(0.026)
College LD Shock	0.014	−0.051	-0.051^{*}	−0.045	-0.046^{**}	-0.071^{**}	-0.059^{*}	-0.055^{*}	−0.047	-0.051^{**}
	(0.023)	(0.026)	(0.024)	(0.024)	(0.017)	(0.023)	(0.027)	(0.024)	(0.024)	(0.017)
Private			−0.001	−0.002	−0.001			−0.001	−0.002	−0.001
			(0.005)	(0.005)	(0.002)			(0.005)	(0.005)	(0.002)
RPU			-0.016^{***}	-0.015^{***}	-0.011^{***}			-0.016^{***}	-0.015^{***}	-0.011^{***}
			(0.005)	(0.004)	(0.003)			(0.005)	(0.004)	(0.003)
Private × In-State			-0.028^{***}	-0.025^{***}	-0.032^{***}			-0.028^{***}	-0.025^{***}	-0.032^{***}
			(0.008)	(0.007)	(0.006)			(0.008)	(0.007)	(0.006)
RPU × In-State			0.014	0.013	0.007			0.014	0.013	0.007
			(0.008)	(0.008)	(0.005)			(0.008)	(0.008)	(0.005)
College In State			0.049^{***}	0.043^{***}	0.050^{***}			0.049^{***}	0.043^{***}	0.050^{***}
			(0.008)	(0.009)	(0.006)			(0.008)	(0.009)	(0.006)
BA by Age 24				0.037^{***}	0.036^{***}				0.037^{***}	0.036^{***}
				(0.005)	(0.004)				(0.005)	(0.004)
Transfer Student				0.003	0.003				0.003	0.003
				(0.002)	(0.002)				(0.002)	(0.002)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Num. obs.	325544	325544	325544	325544	325544	325544	325544	325544	325544	325544

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

First-stage partial F-statistic for high school LD shock is 4795502 ($R^2 = 0.055$) and 6441060 for college shock ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor markets as endogenous independent variables, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) incorporates college graduation cohort fixed effects, Model (3) controls for institutional grouping and residency, and Model (4) adds major, transfer, and on-time completion controls. Model (5) estimates Equation 6. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate resided to the labor market containing their **high school** in the fifth year after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to high school labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. Table 8 reports results of the first-stage estimation process described in Section 4.1.

Table 17: Probability of returning to labor market containing high school and college by 5th year since college graduation

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
HS LD Shock	0.256**	0.013	0.016	0.010	−0.000	0.011	0.028	0.029	0.024	−0.003
	(0.080)	(0.058)	(0.059)	(0.060)	(0.063)	(0.051)	(0.063)	(0.063)	(0.064)	(0.065)
Private, Non-Profit			−0.060	−0.060	−0.067*			−0.060	−0.060	−0.067*
			(0.037)	(0.036)	(0.028)			(0.037)	(0.036)	(0.028)
RPU			−0.069	−0.066	−0.061			−0.069	−0.066	−0.061
			(0.038)	(0.037)	(0.031)			(0.038)	(0.037)	(0.031)
Private × In-State			0.078*	0.076*	0.082**			0.078*	0.076*	0.082**
			(0.039)	(0.037)	(0.029)			(0.039)	(0.037)	(0.029)
RPU × In-State			0.077*	0.076*	0.074*			0.077*	0.076*	0.074*
			(0.039)	(0.038)	(0.031)			(0.039)	(0.038)	(0.031)
College In State			−0.065	−0.062	−0.058*			−0.065	−0.062	−0.058*
			(0.036)	(0.035)	(0.025)			(0.036)	(0.035)	(0.025)
BA by Age 24				0.006	0.007				0.006	0.007
				(0.004)	(0.004)				(0.004)	(0.004)
Transfer Student				0.013***	0.012**				0.013***	0.012**
				(0.004)	(0.004)				(0.004)	(0.004)
Col. Grad. Year FE	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Major FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Transfer FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Group FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Occp. Prestige FE	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
HS State FE	No	No	No	No	Yes	No	No	No	No	Yes
Mean Dep. Var	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13
Num. obs.	59227	59227	59227	59227	59227	59227	59227	59227	59227	59227

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

First-stage partial F-statistic for high school LD shock is 4795502 ($R^2 = 0.055$) and 6441060 for college shock ($R^2 = 0.055$). First-stage F-statistic for high school and college LD shock is 1591231 ($R^2 = 0.055$). The OLS columns use annual employment growth in the high school and college labor market as endogenous independent variables, while the IV columns employ labor demand instrument described in Section 3.1. Models (1) through (4) estimate more parsimonious versions of Equation 6: Model (1) includes no individual controls, Model (2) incorporates college graduation cohort fixed effects, Model (3) controls for institutional grouping and residency, and Model (4) adds major, transfer, and on-time completion controls. Model (5) estimates Equation 6. White standard errors appear in parentheses, clustered at the labor market level. The dependent variable is a binary indicator for whether the college graduate resided to the labor market containing their **high school and college** in the fifth year after college graduation. Coefficients in the first and second rows can be interpreted as semi-elasticities: the percentage point change in the probability of return to high school labor market associated with a 100 percent labor demand shock. The coefficients in the third and fourth rows are relative to Public Flagship graduates. Similarly, the coefficients in the fifth and sixth rows are relative to in-state Public Flagship graduates. Table 8 reports results of the first-stage estimation process described in Section 4.1.

7.4 Descriptives of 1982-2000 Cohorts

Table 18: Summary statistics of excluded cohorts relative to universe

Measure	All Col. Graduates	Full LI Sample	All Grads 1982-99	LI Sample 1982-99
Public Flagship	24	23	25	23
Private, For Profit	2	0	0	0
Private, Non-Profit	32	41	32	41
Regional Public	43	36	43	37
Business Majors	29	21	21	21
Engineering Majors	15	10	6	7
Mobility Rate	2	2	2	2
Mean Parent Income ($)	143232	167552	146020	170583
Mean Kid Income ($)	57045	62675	57845	63453
Average Acceptance Rate	66	63	66	63
90th Percentile	1959	1968	1960	1960
75th Percentile	1968	1980	1963	1964
Median	1979	1991	1968	1969
25th Percentile	1989	1996	1972	1973
10th Percentile	1995	1998	1975	1976
Northeast	22	27	23	24
South	32	28	31	29
Midwest	25	25	27	28
West	20	19	19	17
N	67217457	79852	21945754	15297

Statistics comparing the LI sample to the universe of working college graduates in the 2023 1-year American Community Survey. The first and third columns report ACS data, with the third column restricting to the 1982-1999 college graduation cohorts. The ACS columns condition on labor force participation and a bachelor's degree. The second column characterizes the full LI sample, while the fourth column reports characteristics of LI users who graduated college between 1982-1999. The top four rows in the table report shares of graduates by institutional grouping. The ACS lacks questions about where people obtained their bachelor's degree, so I simulate the distribution of baccalaureate institutions among US workers using institution-level degree awards from 1983-2022 in IPEDS. For pre-1983 graduation cohorts (8 percent of workers in 2023), I used the institutional distribution of degrees awarded in 1983. I exclude Private, For-Profit institutions from the LI sample because most for-profit schools operate online programs, which means the impact on migration and social ties is nebulous. Similarly, I used IPEDS degree awards by year, field, and institution to obtain the business and engineering major shares in the next two rows. The next three rows provide additional comparisons of the institutional distribution, reweighting data from Chetty et al. 2017 and IPEDS based each baccalaureate institution's share. Next, I also compare the age distribution between the LI sample and ACS at the 10th, 25th, 50th, 75th, and 90th percentiles. In the LI data, I estimate year of birth and observe year of college graduation. Thus, I know the joint distribution of birth and college graduation cohorts. Since I only observe year of birth in ACS data, I impute year of college graduation using the joint distribution observed in LI. Finally, I report the share of bachelor's degrees awarded in each Census region, based on the location of the college.

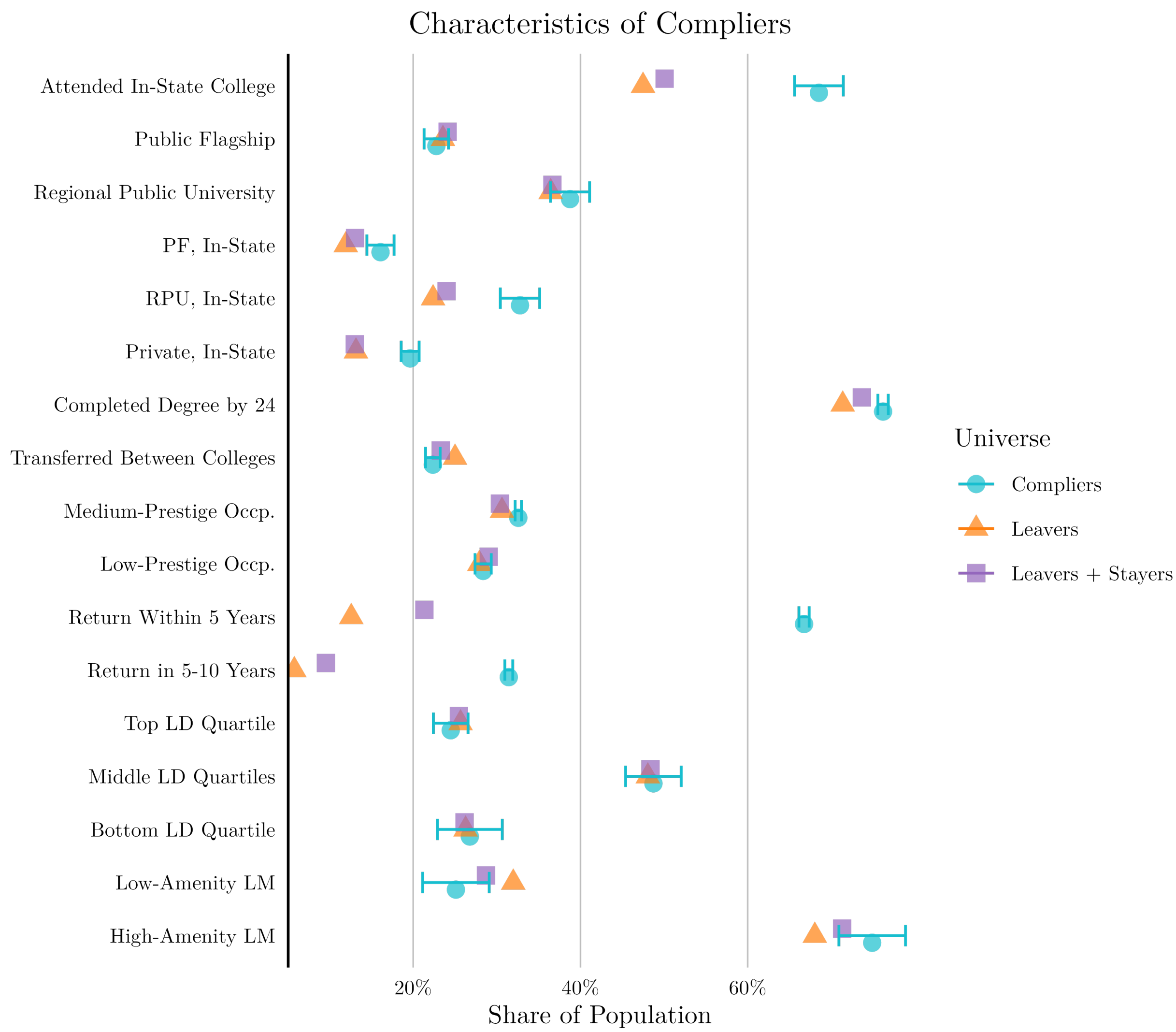

Figure 17: Compliers are college graduates induced to return to their high school labor market by labor demand shocks. Estimation of complier means relies on the Kappa theorem in Abadie 2003. The blue elements with error bars are the complier means and their White standard errors. The orange triangles represent shares for includes all college graduates who moved at some point after graduating college between 1982 and 1999. The purple square includes anybody graduating between 1982 and 1999, regardless of whether they moved after college. Compliers tend to graduate from Regional Public Universities and baccalaureate institutions in the same state as their high schools. In-state graduates are nearly two-thirds of compliers, despite accounting for half of the sample. College graduates who return to origin labor markets mirror the occupational composition of the labor market, with each occupational prestige bin accounting for about one-third of compliers. Table 7 presents the data in tabular form.

8 Theoretical Framework

To understand how individuals make migration decisions, consider how they weigh the costs and benefits of moving (Wozniak 2010; Sprung-Keyser, Hendren, and Porter 2022). Let C be a set of locations, where individual i inhabits one such location j and considers k a potential destination.[31] Individual i belongs to some educational group e, which correlates with human capital. The econometrician cannot directly observe human capital, only its correlates like educational attainment or baccalaureate institutional characteristics (Sprung-Keyser, Hendren, and Porter 2022). Let the model include two time periods: $\{t_0, t_1\} \in T$. The choice problem is the following:

$$\arg\max_{k \in C} \; U(\omega_k) = \left\{ \sum_{t}^{T} \left[\frac{E(w(e)_{kt})}{P_{kt}} - r(e)_{jk} \right] \right\} - o(e)_{jk} \tag{7}$$

where $U(\omega_k)$ is the utility individual i derives from inhabiting location k and e represents an educational group. Education group correlates with human capital. Wozniak 2010 operationalized e as educational attainment. By contrast, I use institutional groupings based on selectivity, flagship status, and control as the basis for e.

Individuals in each educational group expect different wages so $E[w(e)_{kt}]$ is the average wage of education group e in location k at time t. Economic theory predicts a relationship between wages and labor demand shocks: positive demand shocks will boost wages and negative ones will diminish wages (Bartik 1991). P_{kt} is the price level at location k at time t, Thus, $\frac{E(w(e)_{kt})}{P_{kt}}$ measures real wages for educational group e in location k at time t.[32]

Individuals face both recurring and one-time moving costs (Wozniak 2010), represented by $r(e)$ and $o(e)$ respectively. Assume $r(e)_{jk} = o(e)_{jk} = 0$ when $j = k$, so individuals bear no moving costs for remaining in the same location. One-time moving costs consist of expenses typically associated with physical relocation, like beer and pizza for movers, truck rental, and boxes. Recurrent moving costs include the psychological cost of being far from friends and family, return travel costs, and preferences for local amenities in j (Sprung-Keyser, Hendren, and Porter 2022). Thus, the costs depend entirely on subjective considerations like individual preferences towards proximity to family, and local amenities. As a result, measurement of recurrent moving costs isn't possible from existing public data. However, LinkedIn profiles listing high school and college offer a means of (imperfectly) estimating recurrent moving costs. Social ties drive attachment to place, and educational institutions facilitate the formation of social ties (Zabek 2024). If an individual is attached to a

31. For the purposes of my analysis $\{j, k\}$ are sub-state labor markets and C is the United States

32. Yagan 2014 notes individuals move for two reasons. Some migrate for better labor market opportunities, while others migrate to weaker labor markets to enjoy a lower cost of living. By treating relocation decisions as a function of real wages, I implicitly account for spatial variation in the cost of living.

place, the location is likely the site of either secondary or post-secondary educational experiences. Secondary schooling garners particular importance because most people reside with their parents during high school, an effective proxy for familial ties. Thus, I can estimate recurrent moving costs based on the locations of an individual's high school and college. Educational experiences shape the spatial arrangement of an individual's social ties, a substantial component of the recurrent moving costs in migration decisions.

According to the model, individual i seeks to live in a location that maximizes utility relative to all other locations, after accounting for moving costs. Put another way, individuals who relocate must satisfy the inequality:

$$\left\{ \left(\sum_{t}^{T} \left[\frac{E(w(e)_{kt})}{P_{kt}} - r(e)_{jk} \right] - o(e)_{jk} \right) \geq \sum_{t}^{T} \left[\frac{E(w(e)_{jt})}{P_{jt}} \right] \middle| \; j \neq k \right\} \tag{8}$$

For an individual to relocate, their destination must offer a higher wage, even after accounting for fixed and one-time moving costs. The framework helps explain why domestic out-migration is relatively uncommon (Zabek 2024; Hershbein and Stuart 2024). Preferences for one's current location j embedded in $c(e)_{jk}$ raise the cost of moving. As a result, widespread out-migration requires paltry wages or strong attachment to the labor market. High rates of out-migration among low-income workers in Appalachia or the Mississippi Delta satisfy the first condition, driven away due to low equilibrium wages (Sprung-Keyser, Hendren, and Porter 2022). The second condition is the focus of the paper. In general, college graduates migrate at higher rates than their peers without college degrees, suggesting college strengthens attachment to the labor market at the expense of place (Notowidigdo 2020; Wozniak 2010). I want to know whether institutional characteristics like selectivity, flagship status, and control are salient parameters in migration decisions.

Table 19: List of public flagship institutions by state

Institution	State	Institution	State
University of Alaska, Fairbanks	AK	University of Michigan-Ann Arbor	MI
Alabama A&M University	AL	University of Minnesota, Twin Cities	MN
Auburn University	AL	Lincoln University	MO
University of Arkansas	AR	University of Missouri	MO
University of Arkansas at Pine Bluff	AR	Alcorn State University	MS
Arizona State University	AZ	Mississippi State University	MS
University of Arizona	AZ	Montana State University	MT
University of California, Berkeley	CA	North Carolina A&T State University	NC
University of California, Davis	CA	North Carolina State University	NC
University of California, Irvine	CA	University of North Carolina Chapel Hill	NC
University of California, Los Angeles	CA	North Dakota State University	ND
University of California, Merced	CA	University of Nebraska–Lincoln	NE
University of California, Riverside	CA	University of New Hampshire	NH
University of California, San Diego	CA	Rutgers University–New Brunswick	NJ
University of California, San Francisco	CA	New Mexico State University	NM
University of California, Santa Barbara	CA	University of Nevada Reno	NV
University of California, Santa Cruz	CA	Cornell University	NY
Colorado State University	CO	Stony Brook University	NY
University of Colorado Boulder	CO	SUNY Buffalo	NY
University of Connecticut	CT	Central State University	OH
University of the District of Columbia	DC	Ohio State University	OH
Delaware State University	DE	Langston University	OK
University of Delaware	DE	Oklahoma State University	OK
Florida A&M University	FL	Oregon State University	OR
University of Florida	FL	University of Oregon	OR
University of South Florida	FL	Pennsylvania State University	PA
Fort Valley State University	GA	University of Pittsburgh	PA
Georgia Institute of Technology	GA	University of Rhode Island	RI
Unviersity of Georgia	GA	Clemson University	SC
University of Hawaii at Honolulu	HI	South Carolina State University	SC
University of Hawaii at Manoa	HI	South Dakota State University	SD
Iowa State University	IA	Tennessee State University	TN
University of Iowa	IA	University of Tennessee	TN
University of Idaho	ID	Prairie View A&M University	TX
University of Illinois Urbana-Champaign	IL	Texas A&M University	TX
Indiana University Bloomington	IN	University of Texas at Austin	TX
Purdue University	IN	University of Utah	UT
Kansas State University	KS	Utah State University	UT
University of Kansas	KS	University of Virginia	VA
Kentucky State University	KY	Virginia State University	VA
University of Kentucky	KY	Virginia Tech	VA
Louisiana State University	LA	University of Vermont	VT
Southern University and A&M College	LA	University of Washington	WA
University of Massachusetts Amherst	MA	Washington State University	WA
University of Maryland Eastern Shore	MD	University of Wisconsin–Madison	WI
University of Maryland, College Park	MD	West Virginia State University	WV
University of Maine	ME	West Virginia University	WV
Michigan State University	MI	University of Wyoming	WY

The public flagship classifications consists of the union of AAU members and non-tribal Land Grant colleges recognized by the US Department of Education, modifying Bound et al. 2019.