

# The Returns to College(s): Estimating Value-Added and Match Effects in Higher Education\*

Jack Mountjoy<sup>†</sup>

Brent R. Hickman<sup>‡</sup>

July 2020

## Abstract

Students who attend different colleges in the U.S. end up with vastly different economic outcomes. We estimate relative value-added of individual colleges to disentangle causal college impacts from student sorting in producing these outcome disparities. Linking administrative registries of high school records, college applications, admissions decisions, enrollment spells, degree completions, and quarterly earnings spanning the Texas population, we identify college value-added by comparing the outcomes of students who apply to and are admitted by the same set of institutions, as this approach strikingly balances student ability measures across college treatments and renders our extensive set of student covariates irrelevant as controls. We estimate a relatively tight, though non-degenerate, distribution of value-added across the wide diversity of Texas public universities. Selectivity is a poor indicator of value-added, with a negative selectivity effect on STEM major completion and only a fleeting selectivity earnings premium that fades out completely after a few years in the labor market. Non-peer college inputs like instructional spending more strongly predict value-added, especially conditional on selectivity, but residual differences in value-added remain that are not neatly summarized by differences in observable quality measures. Examining potential mechanisms, colleges that ultimately boost earnings also tend to boost persistence, BA completion, and STEM degrees along the way. Finally, we probe the potential for (mis)match effects by allowing value-added to vary flexibly by student characteristics, including race, gender, family income, and pre-college measures of cognitive and non-cognitive skills. At first glance, Black students appear to face small negative returns to attending more selective colleges, but this pattern of modest “mismatch” is driven by the availability of two large historically Black universities with above-average value-added. Across the non-HBCUs, Black students face similar returns to selectivity, and indistinguishable value-added schedules more generally, compared to their peers from other backgrounds.

---

\*We are grateful to Josh Angrist, Raj Chetty, and John Friedman for valuable feedback, along with Dan Black, David Deming, Michael Dinerstein, Peter Ganong, Nathan Hendren, Peter Hull, David Lee, Jordan Matsudaira, Arnaud Maurel, Casey Mulligan, Derek Neal, Chris Neilson, Matt Notowidigdo, Jordan Richmond, Jonah Rockoff, Raffaele Saggio, Owen Zidar, Seth Zimmerman, and seminar participants at Princeton, the Federal Reserve Banks of New York and Chicago, the University of Oslo, Statistics Norway, Duke, Chicago Booth, Chicago Economics, NBER Labor Studies, Washington University in St. Louis, Harvard Opportunity Insights, and Columbia Teachers College. We also thank Rodney Andrews, Greg Branch, Janie Jury, Mark Lu, Greg Phelan, and John Thompson at the UT-Dallas Education Research Center for expert guidance on the administrative data. Nidhaanjit Jain provided excellent research assistance. We are grateful for financial support from the Industrial Relations Section at Princeton University and the Robert H. Topel Faculty Research Fund at the University of Chicago Booth School of Business. The conclusions of this research do not necessarily reflect the opinions or official positions of the Texas Education Research Center, the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas. JEL codes: I21, I24, I26, J31.

<sup>†</sup>Corresponding author. University of Chicago Booth School of Business and NBER. jack.mountjoy@chicagobooth.edu.

<sup>‡</sup>Washington University in St. Louis Olin Business School. hickmanbr@gmail.com.

# 1 Introduction

Marginal students induced into enrolling in college tend to reap significant earnings gains, indicating that college often pays off for students at the outer margins of the higher education system.<sup>1</sup> But for most of the college-bound, the relevant choice is not whether to enroll, but where. Comparing typical student outcomes across institutions suggests this choice may be highly consequential. Graduation rates and mean earnings at the top of the college distribution can be vastly different from those at the bottom: students who attend MIT are nearly four times as likely to graduate (93% vs. 25%) and earn nearly four times as much (\$104,700 vs. \$27,700) as students who attend Alabama State. These divergent outcomes may reflect the impacts of differing college inputs—MIT’s per-student spending on instruction, academic support, and student services in 2018 was nearly \$130,000, while Alabama State’s analogous per-student total was just \$15,000—but little is known about how effectively different colleges use these inputs to create value.<sup>2</sup>

Interpreting differences in student outcomes across colleges is further complicated by the strong sorting of students with vastly different levels of academic preparation and socioeconomic backgrounds into different institutions. To continue the example above, the average incoming student at MIT scored 1530 (out of 1600) on the SAT, while average scores at Alabama State are around 860. Moreover, 74% of students at Alabama State have family incomes low enough to qualify for Pell grants, compared to only 17% at MIT. Different universities likely enroll students with systematically different unobservable traits as well, like ambition and work ethic, that influence a student’s outcomes regardless of where she attends.<sup>3</sup> The selective nature of enrollment leaves us with a difficult question: to what extent do differences in graduation rates and earnings across colleges reveal active institutional impacts on student outcomes, rather than passive reflections of durable student inequalities formed long before enrollment? Understanding how these outcome differences arise is an important consideration for college-bound students deciding where to enroll, as well as for policy-makers deciding where to invest, with state legislators increasingly exploring funding models that directly tie public college appropriations to student outcomes.<sup>4</sup>

In this paper, we estimate relative value-added of individual colleges to isolate causal college impacts from selection bias in producing observed student outcomes. We allow each college to have its own unique

---

<sup>1</sup>E.g. Card (2001); Lemieux and Card (2001); Carneiro et al. (2011); Angrist and Chen (2011); Zimmerman (2014); Heckman et al. (2018); Mountjoy (2019); Smith et al. (2020).

<sup>2</sup>Numbers are taken from U.S. Department of Education sources: graduation rates and earnings from the College Scorecard, <http://collegescorecard.ed.gov>, and per-student spending from IPEDS, <https://nces.ed.gov/ipeds>.

<sup>3</sup>Recent structural evidence by Bodoh-Creed and Hickman (2019) suggests that students’ unobservable characteristics play an important role alongside college quality in driving post-graduate outcomes.

<sup>4</sup>Thirty states base postsecondary appropriations on student outcomes in some way, with at least 11 others in the process of designing such formulas for implementation (National Conference of State Legislatures, 2018).

treatment effect on each outcome, which enables us to explore several interrelated questions. Do some colleges boost student outcomes more than others? Are there observable institutional characteristics, like selectivity or per-student spending, that predict such value-added? Do colleges that boost specific outcomes, like degree completion, also boost others, like earnings? And does value-added vary across students from different backgrounds, producing (mis)match effects?

In answering these questions we must confront three significant challenges. First, estimating unique treatment effects for each college is a data-intensive exercise. In order to learn something meaningful, we need to observe a wide enough swath of colleges to span the quality spectrum, and a large enough sample of students within each college to estimate value-added with meaningful precision. Second, this exercise requires a series of individual-level data linkages, from student characteristics determined prior to college entry, to enrollment spells potentially spanning multiple institutions, to policy-relevant outcomes like degree completions, major choices, and earnings. Third, the college market in the U.S. is a highly decentralized endogenous selection process, both on the part of students deciding where to apply and colleges deciding whom to admit, so identifying causal impacts of colleges requires a research design that credibly reckons with the strong sorting of students with different potential outcomes into different institutions.

We address these challenges by linking multiple administrative registries that span the state of Texas, a setting that offers a vast diversity of colleges and students who attend them. Our data include the entire population of Texas public high school students—nearly 10% of all public school students in the United States—and their college choices across the Texas postsecondary landscape. This solves the first challenge of acquiring large sample sizes of enrollees across a highly heterogeneous array of colleges. We overcome the second challenge by linking the high school records of these students, including rich pre-college measures of skills and backgrounds, to statewide higher education registries that track each student’s college enrollments, degree completions, and major choices, as well as to administrative earnings registries that continue to track these students well into their early careers. To address the third challenge of selection bias, a unique strength of our dataset is the ability to further link these students to administrative records of all applications and admissions decisions at all thirty Texas public universities. Acquiring admissions data from one institution is a difficult task; observing it across the entirety of a very large state’s public university landscape is quite rare, so we focus our analysis on students who begin college at one of these thirty Texas public universities to take advantage of this unique collection of data.

Specifically, we employ the application and admissions records to implement a research design that

compares the outcomes of students who apply to and are admitted by the same set of institutions.<sup>5</sup> The key idea behind this “matched applicant” approach, pioneered by Dale and Krueger (2002) to study linear returns to college selectivity, is that a student’s decisions about which colleges to apply to, and those colleges’ decisions about whether to admit or reject, reveal important information about that student’s abilities, ambitions, and other unobserved advantages that may not be sufficiently captured by typical control variables like test scores and family income.<sup>6</sup> If a student’s portfolio of applications and admissions is a valid proxy for these unobservable factors, such that the remaining variation in potential outcomes is uncorrelated with where students ultimately enroll, then comparing student outcomes within application and admission portfolios identifies the relative value-added of the colleges in students’ choice sets.

We conduct a battery of exercises to probe the validity of this approach in recovering causal comparisons. To address the concern that students facing the same choice set of college acceptances will further sort into institutions systematically by ability, we find that controlling solely for admission portfolio fixed effects strikingly balances student ability measures across college treatments: students who apply to and are admitted by the same institutions have very similar 10th grade test scores regardless of where they actually enroll, despite massive raw differences in scores across campuses. More generally, we show that controlling solely for admission portfolios leaves little remaining role for selection bias driven by the types of observables common to many educational datasets, like test scores and demographics. Specifically, including or excluding these extra covariates makes very little difference in the college value-added estimates after controlling solely for admission portfolio fixed effects. Perhaps more strikingly, we also add several less commonly available controls—including additional measures of academic preparation like advanced high school coursework, behavioral measures of non-cognitive skills like attendance and disciplinary records, and fixed effects for every Texas high school to flexibly account for the environmental influences of schools, neighborhoods, and local labor markets—and again we find little evidence of omitted variable bias, despite the relevance of all of these covariates in predicting student outcomes, even within admissions sets.

Conversely, if we begin with the typical selection-on-observables approach by controlling for commonly

---

<sup>5</sup>Many of the colleges in our data and across the U.S. practice “holistic admissions” with no sharp quantitative admissions cutoffs. This hampers a regression discontinuity (RD) research design, which has been increasingly employed to study admission impacts in countries with centralized college admissions systems (e.g. Saavedra, 2009; Hastings et al., 2014; Kirkeboen et al., 2016; Canaan and Mouganie, 2018; Anelli, 2019; Jia and Li, 2019; Sekhri, 2020). Because the U.S. has no such centralization, the handful of RD studies in the U.S. focus on binary treatment comparisons of getting marginally admitted to a specific institution (Hoekstra, 2009; Zimmerman, 2014), or a set of cutoff-sharing institutions (Goodman et al., 2017), relative to a composite counterfactual of various fallback options. Such a composite counterfactual pools other 4-year institutions, 2-year colleges, and non-attendance together into a single comparison group, hampering the ability of the RD approach to identify college-by-college value-added in the decentralized U.S. higher education system.

<sup>6</sup>Subsequent variants of the matched applicant approach include Arcidiacono (2005), Long (2008), Fryer and Greenstone (2010), Broecke (2012), Cunha and Miller (2014), Arcidiacono et al. (2016), Ge et al. (2018), Eller (2019), and Abdulkadiroglu et al. (2020).

available covariates, and even adding the less commonly available ones, the subsequent inclusion of our admission portfolio fixed effects does appear to account for substantial residual selection. Intuitively, the admission portfolios are so informationally rich because they allow each student to provide a high-dimensional signal of her abilities, ambitions, and advantages by assembling a personalized list of college applications, and they allow the admissions officers of those colleges to contribute an additional high-dimensional signal of their assessment of the student through their collective admissions decisions. Ultimately, we cannot verify that all possible sources of omitted variable bias are eliminated by our matched applicant approach. However, any remaining bias would have to occur on dimensions of unobserved ability or advantage that are not well proxied by our extensive set of observable student covariates, which are themselves rendered unnecessary as controls after conditioning solely on admission portfolios.<sup>7</sup>

Our results reveal several notable patterns in the distribution of value-added across the wide diversity of Texas public universities. First, in terms of magnitude, we estimate a relatively tight, though non-degenerate, distribution of college value-added: moving up one standard deviation in the value-added distribution leads to a 3.7 percentage point gain in the probability of completing a BA, and a \$1,300 boost in earnings around age 28, roughly a 3% premium. Value-added thus varies non-trivially across colleges, though these causal differences are dwarfed by the variation in raw outcome means, which have a standard deviation of 18 percentage points for BA completion and \$8,000 for earnings. A college’s raw BA completion rate is decently correlated with its value-added on BA completion, but this is much less true for earnings: raw earnings means and earnings value-added have a correlation of only 0.2, even after correcting for minor attenuation from estimation error. These results suggest that both the ordering and the magnitudes of the raw outcome comparisons often highlighted in college guides, the popular press, and state funding formulas are driven largely by selection bias from student sorting, leaving a relatively smaller role for causal contributions of the colleges in students’ admission sets.

Relatedly, while college selectivity is a nearly perfect predictor of raw outcome means across colleges, it is a poor indicator of value-added. Students who attend the more selective colleges in their admission portfolios are only slightly more likely to complete a BA, and are significantly less likely to successfully major in STEM—a result that strongly switches signs when moving from raw correlations to causal value-added. With respect to earnings, we do find a modest selectivity premium in a student’s first few years in the

---

<sup>7</sup>A more subtle concern is that different students may face different value-added schedules across colleges and may make enrollment decisions based in part on these idiosyncratic returns. In Section 3 we derive conditions under which we recover the mean value-added schedule across students using our baseline empirical specification, which follows the value-added literatures on teachers (e.g. Chetty et al., 2014c), schools (e.g. Angrist et al., 2017), firms (e.g. Abowd et al., 1999), neighborhoods (e.g. Chetty and Hendren, 2018), etc. Furthermore, we relax this canonical specification in Section 8 and directly investigate the presence of match effects between students and colleges.

post-college labor market, but this premium fades out quickly to a stable and precise zero return thereafter, suggesting that students with similar potential who choose to attend less selective colleges catch up to their elite peers in a labor market that learns about productivity relatively quickly. Moving beyond selectivity, non-peer college inputs like instructional expenditures per student and the share of faculty who are full-time do covary positively with value-added, especially conditional on selectivity, though residual differences in value-added remain that are not neatly summarized by differences in observable quality measures. We also explore the potential for college-level intergenerational mobility statistics to serve as predictors of causal value-added; we do not find strong correlations between these measures, perhaps because mobility statistics that condition solely on family income remain at risk of reflecting other systematic student differences across campuses, like levels of academic preparation and ambition, rather than causal college impacts. Taken together, this set of results suggests that various prominent measures of college “quality” are neither interchangeable nor sufficient as indicators of causal value-added on student outcomes.

Turning to mechanisms, we find that our value-added estimates for each college tend to be correlated across different student outcomes, shedding light on the potential channels through which colleges shape the path from enrollment to earnings. Colleges that boost BA completion also tend to boost earnings: a 10 percentage point increase in BA value-added is associated with a roughly \$3,000 increase in earnings value-added across colleges. This relationship strengthens considerably as we lengthen the degree completion window from 4 years to 6 years to 8 years, indicating that strict measures of on-time graduation may underestimate a college’s ultimate value-added in the labor market. Unpacking BA completion by major category, we find that a college’s value-added on the probability of completing a STEM major has a strong correlation with its earnings value-added, while value-added on completing non-STEM majors has almost no bivariate relationship with earnings value-added. A college’s STEM and non-STEM value-added are negatively correlated, however; conditional on STEM value-added, the relationship between value-added on non-STEM degrees and earnings becomes reasonably positive, though still substantially weaker than the STEM-earnings relationship. We also find predictive power in value-added on non-degree outcomes like persistence and transfer, as well as on industry of employment, highlighting the multiple pathways along which colleges shape student outcomes.

Finally, we probe the potential for (mis)match effects by allowing college value-added to vary flexibly by student characteristics, including race, gender, family income, and pre-college measures of cognitive and non-cognitive skills. Overall, we find little evidence of significant heterogeneity in value-added along most of these dimensions of student diversity: estimates by student subgroups tend to cluster around our main

pooled estimates, and interaction effects between college indicators and student characteristics have very little explanatory power on outcomes. At first glance, one modest exception is that Black students appear to face small negative impacts of attending more selective colleges, but this minor pattern of “mismatch” is driven by the presence of two large historically Black universities in Texas that have low average incoming SAT scores but yield above-average value-added. These HBCUs are over 95 percent Black, and thus do not tend to enter the choice sets of non-Black students. This racial heterogeneity in choice sets, rather than mismatch per se, drives the result, as Black students experience similar impacts of selectivity, and value-added schedules more generally, across the non-HBCUs compared to their peers from other backgrounds.

Our methods and results advance several strands of the literature on the consequences of college choices. First, a large body of prior work has sought to estimate the return to selectivity, or college “quality” more generally, typically by specifying earnings as a function of the mean SAT score of a student’s college peers.<sup>8</sup> While several papers in this strand of literature have estimated significant gains to attending a more selective college, Dale and Krueger (2002, 2014) stand out as an exception after accounting for student application behavior. On one hand, we replicate, validate, and extend this null return to selectivity (after documenting a fleeting initial premium) in our administrative data, which span a much larger, more diverse, more recent, and more precisely measured sample of students and colleges than the College and Beyond survey at the center of Dale and Krueger (2002, 2014)’s analysis. On the other hand, our results on non-peer college inputs build on the work of Black and Smith (2006) and others in exploring other dimensions of college quality that may not be captured by selectivity. Our work thus sheds light on a long-standing apparent tension in this literature: while college selectivity—which, by definition, largely reflects endogenous selection of students with different abilities into different colleges—fails to predict causal value-added within admission sets, non-peer college inputs—which are correlated with selectivity, but not perfectly so—do appear to boost student outcomes, or at least correlate with factors that do.

More importantly, our college-by-college value-added approach allows us to move beyond analyses of college “quality” that confine college heterogeneity to a pre-specified observable dimension, like selectivity or spending. Analogous to the evolution of the teacher value-added literature, by flexibly allowing each college to impart its own unique impact on student outcomes, we expose differences in value-added across

---

<sup>8</sup>See Brewer and Ehrenberg (1996) for a review of the early literature. Black and Smith (2006) and Dillon and Smith (2018) employ a more comprehensive definition of college quality as an index of multiple institutional measures. Bodoh-Creed and Hickman (2019) endogenize pre-college human capital investment in a structural analysis of college sorting and student outcomes. Ge et al. (2018) expand the sample and outcomes of Dale and Krueger (2002) to study impacts of selectivity on female labor supply and family outcomes. Other related work examines impacts of attending different categories of institutions, e.g. flagship vs. non-flagship public universities (Andrews et al., 2016; Black et al., 2020), 4-year vs. 2-year colleges (Reynolds, 2012; Mountjoy, 2019), and for-profit vs. public institutions (Armona et al., 2018; Cellini and Turner, 2019).

colleges that are not neatly summarized by differences in their observable quality measures. Our work thus contributes to a nascent set of papers that have begun to leverage large administrative data sources in the U.S. to examine student outcomes at the granularity of individual colleges, including Cunha and Miller (2014), Arcidiacono et al. (2016), Hoxby (2019), and Chetty et al. (2020). We advance this recent work by merging in administrative admissions records that permit a precise implementation of the matched applicant approach, and we conduct a battery of validation exercises to probe the ability of this approach to disentangle causal college contributions from selection bias.<sup>9</sup> We also estimate college value-added on a significantly expanded suite of student outcomes—including persistence, transfer, degree completion and its timing, major choice, migration, industry of employment, and earnings dynamics—illustrating the variety of interconnected channels through which colleges can influence student trajectories.

Furthermore, our ability to merge in a rich set of pre-college measures of student skills and backgrounds allows us to explore heterogeneity in college value-added schedules faced by different types of students, and bring these estimates to bear on related academic and policy debates (e.g. Bowen and Bok, 2000; Sander and Taylor, 2012; Arcidiacono and Lovenheim, 2016). For example, critics of race-based affirmative action often argue that it harms its intended beneficiaries through a mismatch effect, whereby minority students who attend selective colleges would have fared better in terms of graduation rates and STEM degree completion by choosing less selective schools instead.<sup>10</sup> We contribute a suite of additional policy-relevant outcomes, including earnings, to inform this and other related questions about heterogeneous impacts, including whether low-income students benefit more or less than their higher-income peers from choosing selective colleges, and whether a student’s own abilities interact with her institution to generate complementarities in the production of value-added. Our results indicate only a limited amount of heterogeneity in college value-added across these observable dimensions of student diversity, with the slight apparent mismatch effect for Black students driven by heterogeneity in alternatives (i.e. the availability of HBCUs) rather than significant heterogeneity in the value-added reaped by different students at any given university.<sup>11</sup>

In broadly interpreting our results, it is important to keep in mind that our matched applicant research design, like most quasi-experimental observational studies, may sacrifice some external validity on the al-

---

<sup>9</sup>Cunha and Miller (2014) offer a pioneering proof-of-concept demonstration of the use of admission portfolio fixed effects in estimating college value-added in Texas. For computational reasons, they grouped all students who applied to four or more schools into a single portfolio control cell; their sample also included students with no application or admissions data, grouped into another single control cell. We find that these aggregations appear to substantially under-control for selection, as these two coarse control cells comprise one fourth of their identifying sample (students in control cells with treatment variation) and permit significant residual sorting into college treatments, given the uncontrolled heterogeneity of the students within them.

<sup>10</sup>For prominent legal articulations of this view, see the opinion of Clarence Thomas in *Fisher v. University of Texas at Austin*, 570 U.S. 297 (2013), and oral arguments by Antonin Scalia in *Fisher v. University of Texas at Austin*, 579 U.S. (2016).

<sup>11</sup>This result is similar in spirit to Angrist et al. (2019), who explore the roles of match effects versus treatment alternatives in explaining the impacts of selective high school admissions on student outcomes.



tar of internal validity, in our case by honing the scope of identifying variation to college choices within admission sets. Our variation is highly relevant for learning about the long-term consequences of choices made after multiple admission letters arrive. This variation does not necessarily speak to impacts at other important margins upstream, like admitting minimally qualified applicants just above a cutoff (e.g. Hoekstra, 2009; Zimmerman, 2014; Goodman et al., 2017), or intervening early in the application process to transform disadvantaged students’ fundamental college ambitions (e.g. Hoxby and Turner, 2013; Dynarski et al., 2020).<sup>12</sup> Research designs that employ variation at these other margins not only involve different student subpopulations with potentially different treatment responses, but also involve substantively different counterfactuals, usually pooling other 4-year institutions, 2-year colleges, and non-enrollment into a composite comparison group. Our college-by-college value-added estimates for inframarginal admitted students thus complement results from these other research designs, hopefully contributing to a more comprehensive understanding of causal consequences at distinct junctures in the higher education pipeline.

The paper proceeds as follows. Section 2 describes and summarizes our data on students and colleges. Section 3 defines our value-added target parameters in a potential outcomes framework and discusses how the matched applicant approach addresses threats to their identification. Section 4 presents the main value-added estimates and conducts a battery of validation exercises and robustness checks. Section 5 quantifies distributional magnitudes of value-added across colleges. Section 6 investigates institutional predictors of value-added. Section 7 explores value-added on intermediate outcomes as potential mechanisms driving earnings effects. Section 8 investigates heterogeneity in value-added via match effects. Section 9 concludes.

## 2 Data

In this section, we describe our setting and data sources. We also define our analysis sample and variables of interest, and we present summary statistics on students and colleges.

### 2.1 Setting, Sample, and Sources

Our data come from linking several administrative registries that span the entire state of Texas. Texas is the second largest U.S. state by population, land area, and GDP, with nearly 30 million residents and an economy that would rank 10th largest in the world as a sovereign nation. This immensity supports a comprehensive higher education system that enrolls 1.6 million students and provides a fruitful setting for

---

<sup>12</sup>Our variation is also limited to in-state public institutions. Most students in the U.S. attend such colleges, and the Texas public higher education market is one of the largest and most diverse in the country, but our variation may not speak to comparisons that involve significantly more expensive and well-resourced private colleges (e.g. Cohodes and Goodman, 2014).

studying value-added across a rich diversity of colleges and students.

Our analysis sample begins with the population of students who graduated from a Texas public high school between 1999 and 2008.<sup>13</sup> We link several student registries maintained by the Texas Education Agency (TEA) to assemble pre-college data on these students' demographics, standardized test scores, academic preparation, attendance, disciplinary infractions, and school/neighborhood environment. We then link these high school graduates to administrative application and admissions records from all Texas public 4-year universities, maintained by the Texas Higher Education Coordinating Board (THECB). To take advantage of this uniquely comprehensive repository of admissions data, we focus our analysis sample on students who begin college at one of these 30 Texas public universities.<sup>14</sup> We then follow these students longitudinally through administrative registries of all college enrollment spells and degree completions (also using THECB data) spanning all public and private non-profit postsecondary institutions in the state, allowing us to observe educational outcomes inclusive of transfer. Finally, we merge in quarterly earnings records for these students from the Texas Workforce Commission (TWC) that cover all Texas employees subject to the state unemployment insurance system.<sup>15</sup> We supplement the student-level data with college-level institutional characteristics from the Integrated Postsecondary Education Data System (IPEDS) and college-level intergenerational mobility statistics from Chetty et al. (2020).

As with any data spanning a particular state, we are unable to track activity outside of Texas. In our case, this means we will not observe college enrollments, degree attainments, and earnings for students who outmigrate. Fortunately in our case, Texas has the lowest outmigration rate of any state in the U.S. (Aisch et al., 2014). Auxiliary data from the National Student Clearinghouse, which tracks nationwide college enrollments, show that less than 4 percent of the 2008 and 2009 cohorts of Texas public high school graduates left the state to attend college (Mountjoy, 2019). With respect to earnings, we are unable to distinguish nonemployment within Texas from unobserved employment outside the state, so we condition our labor market outcomes on having at least some positive Texas earnings. We conduct a battery of robustness checks in Section 4.3 and fail to find any evidence that endogenous selection out of the earnings sample introduces bias in our value-added estimates.

---

<sup>13</sup>The 1999 cohort is the first to have college application and admission data, and we stop at the 2008 cohort to observe at least 10 years of post-high-school earnings for all sample members. Private high school students are not observed in this data; they account for only 5 percent of all Texas high school graduates (National Center for Education Statistics, 2018).

<sup>14</sup>We do not observe applications or admissions records at private Texas colleges. Our 30 public universities enroll roughly 86% of 4-year college-goers in Texas. This set also excludes small and young public institutions with limited or no first-year enrollment during our sample period. We do not observe enrollments or degrees at for-profit colleges, which accounted for less than 5 percent of national enrollments during our sample period (National Center for Education Statistics, 2018).

<sup>15</sup>Stevens (2007) estimates that 90% of the civilian labor force is captured in state UI records; excluded are the self-employed, independent contractors, some federal employees including military personnel, and workers in the informal sector.

## 2.2 Variable Definitions

*Pre-College Covariates*—Our pre-college student covariates from the TEA high school files fall into four categories: demographics, test scores, other measures of academic preparation, and behavioral/disciplinary records. For demographics, we observe each student’s gender, race, and free or reduced-price lunch eligibility as a proxy for low family income.<sup>16</sup> We also observe unique indicators for each high school, which we use in robustness checks to flexibly capture environmental influences of the schools, neighborhoods, and local labor markets that students experience prior to college entry. For standardized test scores we measure achievement on mandatory statewide 10th grade math and reading exams by standardizing the raw scores to have mean zero and standard deviation one within each cohort. We also observe SAT scores from the College Board for the subsample of students who graduate high school in 2003-2008. Other measures of academic preparation include the number of Advanced Placement, International Baccalaureate, and other advanced courses taken, as well as indicators for whether a student was ever at risk of dropping out of high school and (from the THECB admissions data) whether she was in the top 10% of her high school class. Finally, we measure attendance behavior as the share of total school days in 10th-12th grade that a student is not absent, and we measure disciplinary infractions as the average number of days suspended in 10th-12th grade, multiplied by  $-1$  to covary positively with academic and labor market outcomes.

*Application/Admission Portfolios*—To compare students who apply to and are admitted by the same institutions, we construct unique indicators for each distinct portfolio of observed college applications and admissions decisions. With respect to each college there is a ternary outcome—*do not apply*, *apply and get rejected*, or *apply and get admitted*—and there are 30 colleges, so theoretically there are  $3^{30}$  possible unique portfolios. In practice, we observe 21,332 distinct portfolios in our data, since most students only apply to a small number of schools. We show in Section 4.2 that we can reduce the dimensionality of these portfolios without affecting our value-added estimates by discarding information about rejections; that is, by reducing the college-specific outcome to a binary of *do not get admitted* vs. *apply and get admitted*, where *do not get admitted* could be due either to not applying or applying and getting rejected.<sup>17</sup> There are 8,002 such distinct portfolios of admissions to our 30 institutions, of which 4,689 are unique to a single student and thus do not contribute to identification in our portfolio fixed effects approach. The remaining

---

<sup>16</sup>The three race categories are white, Black, and Hispanic. Asians are pooled with whites and Native Americans are pooled with Hispanics due to small shares at several colleges.

<sup>17</sup>The binary portfolio specification effectively condenses multiple ternary portfolios into a single portfolio. For example, suppose Tom applies to colleges A, B, and C, and is admitted by A and C, while Harry applies only to A and C, and is admitted by both. The ternary specification will consider these as two separate portfolios, while the binary specification considers them as the same one. In Section 4.2 we show that both methods yield nearly identical estimates, suggesting that, conditional on colleges’ collective acceptance decisions, rejections contain no additional identifying information.

3,313 admissions portfolios with at least two students form the basis of our research design, which compares outcomes of students who reveal themselves to be sufficiently qualified to get admitted to the same set of schools but end up making different enrollment choices.

*Treatments*—We allow each of the 30 public universities in Texas to have a unique treatment effect (value-added) on student outcomes. We operationalize this by constructing mutually exclusive and exhaustive treatment indicators for each school, with UT-Austin omitted as the comparison institution against which all the others are measured. We define each student’s treatment using their first college enrollment starting in the fall after high school graduation.<sup>18</sup> This framework fully permits transfer and drop-out behavior by classifying these subsequent events as endogenous intermediate outcomes, potentially influenced by the initial college treatment. That is, the choice of where to *begin* college may well affect how long a student stays enrolled and whether she transfers to additional institutions, perhaps even for longer durations than the initial spell. See Hoxby (2019) for an insightful discussion of this downstream “attribution problem” in assessing the impacts of colleges on ultimate student outcomes.

*Academic Outcomes*—We study causal value-added of colleges on several academic outcomes along the path from a student’s first college enrollment to her last. To study endogenous transfer behavior, we construct an indicator for whether a student subsequently enrolls in a different institution from her initial treatment.<sup>19</sup> We measure BA completion as an indicator for obtaining a baccalaureate degree from any public or private institution in the THECB registries, which avoids the penalization of transfer students that commonly occurs in datasets that can only track students within their initial institution of enrollment. We also capture dynamics in time-to-degree by constructing separate indicators for completing within a certain number of years: 4, 6, or 8. For a more continuous measure of persistence, we also calculate years of college completed.<sup>20</sup> Finally, we observe detailed major (field of study) codes for students who complete a BA, so analogously to the BA completion outcome we construct separate indicators for completing a STEM (science, technology, engineering, or mathematics) degree within 4, 6, and 8 years.

*Labor Market Outcomes*—Our main labor market outcome is annualized earnings around ten years out from college entry, which we measure as follows. We begin by summing earnings within each person-quarter across different jobs and deflating by the quarterly U.S. consumer price index (base year 2015).

---

<sup>18</sup>We follow Andrews et al. (2014) and ignore summer terms when defining the institution of first enrollment, as some students take supplemental summer courses at local campuses before starting their main degree program.

<sup>19</sup>To avoid misclassifying supplemental coursework and dual enrollment, we ignore transfer episodes in which the student returns to the initial institution.

<sup>20</sup>Specifically, years of college completed in our college-going sample range from 0 to 4. To be classified as completing 0 years: complete high school but less than one year of college. To complete 1: enroll in college with 2nd year standing, or complete a certificate, or complete the academic core requirement at a community college. 2: enroll in college with 3rd year standing, or complete an associate’s degree. 3: enroll in college with 4th year standing. 4: complete a bachelor’s degree.

We winsorize at the top and bottom 0.1<sup>th</sup> percentiles and average the non-missing earnings quarters within person over the range of 8-10 years out from college entry, roughly ages 27-29, which are the oldest ages that all of our analysis cohorts have in common.<sup>21</sup> We arrive at an annualized measure by multiplying this quarterly average by 4. We also construct a panel of person-year earnings to explore earnings trajectories. Since the TWC data contain industry of employment, we also construct an indicator for working in the oil and gas industry to explore this as a potential mediator of colleges' earnings impacts. Finally, to probe the robustness of our earnings value-added estimates to sample selection from outmigration or nonemployment, we construct an indicator for simply appearing in the TWC earnings data during our measurement ages and conduct checks with this sample indicator as an intermediate outcome.

### 2.3 Summary Statistics

Table 1 presents summary statistics for our analysis sample. Our ten cohorts of Texas public high school graduates who begin college at one of the 30 Texas public universities comprise 422,956 students, of whom 54 percent are female, 24 percent are low-income (eligible for free or reduced-price lunch in high school), 12 percent are Black, and 23 percent are Hispanic. 60 percent of students only apply to one Texas public university, and 69 percent are only admitted to the school they end up attending. These students do not contribute to the identification of college value-added in our matched applicant approach, since there is no variation in treatment within their single-college admission portfolios.

In terms of treatments, roughly one quarter of our sample begin college at one of the two flagships, UT-Austin and Texas A&M, which are among the largest universities in the country. The University of Texas (UT) System included 7 other standalone institutions during our sample period (UT-San Antonio, UT-Pan American, UT-Arlington, UT-El Paso, UT-Dallas, UT-Tyler, and UT-Permian Basin), and the Texas A&M (TAMU) System included 8 other standalone institutions (Tarleton State, TAMU-Corpus Christi, Prairie View A&M, TAMU-Kingsville, West Texas A&M, TAMU-Commerce, TAMU-International, and TAMU-Galveston). Three other public systems also oversee multiple standalone institutions: the Texas State System (Texas State-San Marcos, Sam Houston State, Lamar, and Sul Ross State), the Texas Tech System (Texas Tech and Angelo State), and the University of Houston System (Houston and Houston-Downtown). The University of North Texas belongs to its own system, and the remaining universities—Stephen F. Austin State, Midwestern State, Texas Southern, and Texas Woman's—operate independently

---

<sup>21</sup>We show below in Section 4.3 that our main value-added estimates are highly correlated with estimates from the subsample of older cohorts for whom we can measure earnings at later ages, and in Section 6 we document that the earnings effects of attending a more selective college flatten out by 8-10 years out.

Table 1: Summary Statistics

	Mean	(SD)		Count	Share
<i>Covariates</i>			<i>Treatments</i>		
Female	.544		Texas A&M (TAMU)	54,953	.13
Low-income (FRPL)	.241		UT-Austin	52,508	.124
Black	.121		Texas Tech	32,371	.077
Hispanic	.227		UT-San Antonio	27,569	.065
10th grade test score (std.)	0	(1)	North Texas	24,146	.057
High school attendance (std.)	0	(1)	Texas State-San Marcos	23,686	.056
			Houston	23,528	.056
<i>Applications</i>			Stephen F. Austin State	17,372	.041
Applied to 1 school	.601		Sam Houston State	15,704	.037
Applied to 2 schools	.233		UT-Pan American	15,000	.035
Applied to 3 schools	.104		UT-Arlington	14,595	.035
Applied to 4 schools	.041		UT-El Paso	14,361	.034
Applied to 5+ schools	.022		Angelo State	10,585	.025
			Lamar	10,569	.025
<i>Admissions</i>			Tarleton State	9,795	.023
Admitted to 1 school	.691		TAMU-Corpus Christi	8,550	.02
Admitted to 2 schools	.212		Texas Southern	7,736	.018
Admitted to 3 schools	.069		Prairie View A&M	7,353	.017
Admitted to 4 schools	.02		TAMU-Kingsville	6,675	.016
Admitted to 5+ schools	.009		West Texas A&M	6,498	.015
			UT-Dallas	6,453	.015
<i>Academic Outcomes</i>			Midwestern State	5,873	.014
Ever transfer	.271		Houston-Downtown	5,196	.012
Years of college completed	2.89	(1.52)	TAMU-Commerce	4,293	.01
BA within 4 years	.274		Texas Woman's	4,001	.009
BA within 6 years	.592		TAMU-International	3,537	.008
BA within 8 years	.652		UT-Tyler	3,248	.008
STEM degree	.13		TAMU-Galveston	2,797	.007
			Sul Ross State	2,037	.005
<i>Earnings Outcomes</i>			UT-Permian Basin	1,963	.005
Has positive earnings	.848				
Annualized earnings	44,834	(28,485)			
			Observations	422,956	

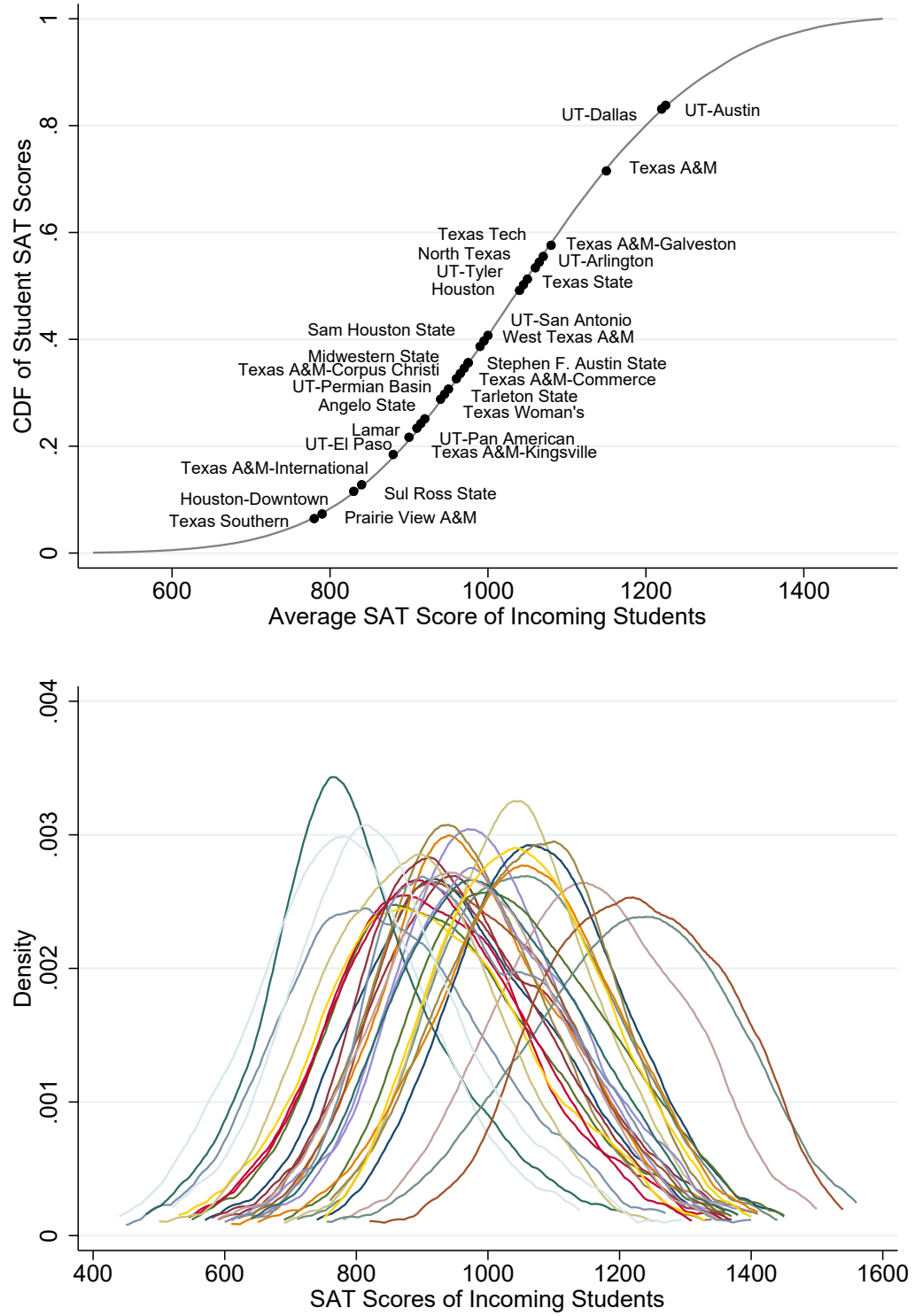
Notes: The sample is comprised of students who graduate from a Texas public high school in 1999-2008 and enroll at one of the Texas public universities listed in the Treatments column. Treatments are defined as the first college attended in the school year after high school graduation. *FRPL* denotes free or reduced price lunch eligibility in high school. *10th grade test score* is an average of math and reading scores and *High school attendance* is measured as days not absent in 10th-12th grade; both of these are standardized to have mean 0 and standard deviation 1. Academic outcomes are measured at 6 years from college entry unless otherwise noted. Earnings outcomes are measured 8-10 years from college entry in real 2015 dollars.

of any system.

The top panel of Figure 1 illustrates the vast diversity of these institutions in terms of selectivity. Each school's mean SAT score of incoming students is plotted on the student-level CDF of SAT scores, showing that our campus averages span nearly 80 percentiles of the student-level score distribution.<sup>22</sup> Despite these large cross-campus differences in means, the bottom panel of Figure 1 shows that the majority of the student-level variation in SAT scores is actually within campuses; a variance decomposition attributes 75% within campuses and 25% across campuses. In our case, this is an attractive feature of the Texas

<sup>22</sup>Coincidentally, while we do not observe schools at the very top of the selectivity distribution, this missing complement of our data is precisely the range of selectivity that Dale and Krueger (2002, 2014) focus on in their main dataset, the College and Beyond survey. Our results thus complement and extend their work by studying the lower 80 percent of the selectivity distribution not captured in the highly selective schools of the C&B survey.

Figure 1: Selectivity Across Colleges and SAT Score Variation Within Colleges



Notes: The top panel positions the average SAT score of incoming students at each college in our sample within the overall CDF of student-level SAT scores. The bottom panel plots the distribution of SAT scores within each college.

data, as the overlapping supports provide the kind of within-institution variation necessary for identifying (mis)match effects across heterogeneous students.

Finally, in terms of outcomes, Table 1 shows that beginning college is far from synonymous with finishing a degree, especially on time: only 27% of students complete a BA within 4 years of entry, with completion rates climbing to 59% and 65% by 6 years and 8 years, respectively (in line with national statistics). On average, our matriculants complete 2.9 years of college, and 27% transfer to another Texas institution (including private colleges and 2-year community colleges) at least once.<sup>23</sup> Thirteen percent of students successfully complete a STEM degree, such that 22% (.13/.59) of all BA completions are in STEM fields. Finally, 85% of our students appear in the earnings data 8-10 years out from college entry, with real annualized earnings averaging around \$45,000 with a standard deviation of nearly \$30,000.

### 3 Research Design

In this section, we define our parameters of interest—causal value-added of individual colleges on student outcomes—in a potential outcomes framework, which clarifies the threats to their identification and the assumptions under which our matched applicant approach addresses these threats.

#### 3.1 Potential Outcomes Framework

Let  $Y_{ij}$  denote the potential outcome of student  $i \in \mathcal{I}$  if she were to enroll at counterfactual college  $j \in \mathcal{J}$ . We can decompose  $Y_{ij}$  as the sum of a constant, a school fixed effect, a student fixed effect, and a residual:

$$\begin{array}{ccccccc}
 Y_{ij} & = & \underbrace{\kappa} & + & \underbrace{\nu_j} & + & \underbrace{\alpha_i} & + & \underbrace{\epsilon_{ij}} & (1) \\
 & & \mathbb{E}[Y_{ij}] & & \mathbb{E}_{\mathcal{I}}[Y_{ij}] - \kappa & & \mathbb{E}_{\mathcal{J}}[Y_{ij}] - \kappa & & Y_{ij} - \kappa - \nu_j - \alpha_i \\
 & & \text{Constant} & & \text{School fixed effect} & & \text{Student fixed effect} & & \text{Residual} \\
 & & \text{(Normalization)} & & \text{(Value-added)} & & \text{(Ability, ambition, advantage)} & & \text{(Idiosyncratic match)}
 \end{array}$$

where  $\mathbb{E}_{\mathcal{I}}$  denotes taking the expectation across the population of students while fixing school  $j$ , and  $\mathbb{E}_{\mathcal{J}}$  takes the expectation across all schools while fixing student  $i$ . Each potential outcome  $Y_{ij}$  begins with the constant  $\kappa$ , the unconditional mean potential outcome. Each school then contributes a fixed value-added parameter  $\nu_j$ , which is the average treatment effect of attending school  $j$  relative to the average school.

<sup>23</sup>Auxiliary data from the National Student Clearinghouse, available for students who began at Texas public colleges in 2011 and 2012, show that only 2.3% transfer out-of-state, and less than 1% earn an out-of-state degree, so our within-Texas estimates are unlikely to be meaningfully biased by out-of-state transfers.



Each student, meanwhile, contributes a fixed effect  $\alpha_i$ , which is her average potential outcome across all schools relative to the average student. This represents vertical traits like ability, ambition, and advantage that boost student  $i$ 's outcomes regardless of where she enrolls. Finally,  $\epsilon_{ij}$  captures any match effects and idiosyncratic shocks not explained by the additively separable school and student fixed effects.

Let  $D_{ij}$  denote a dummy variable indicating whether student  $i$  actually enrolls in school  $j$ , and let  $j = 0$  denote the omitted reference school. We can then write the observed outcome for student  $i$  as

$$\begin{aligned}
Y_i &= Y_{i0} + \sum_{j \neq 0} (Y_{ij} - Y_{i0}) D_{ij} \\
&= \kappa + \nu_0 + \alpha_i + \epsilon_{i0} + \sum_{j \neq 0} (\nu_j + \epsilon_{ij} - \nu_0 - \epsilon_{i0}) D_{ij} \\
&= \tilde{\kappa} + \sum_{j \neq 0} \tilde{\nu}_j D_{ij} + \alpha_i + \tilde{\epsilon}_i,
\end{aligned} \tag{2}$$

where  $\tilde{\kappa} \equiv \kappa + \nu_0$ ,  $\tilde{\nu}_j \equiv \nu_j - \nu_0$ , and  $\tilde{\epsilon}_i \equiv \sum_j \epsilon_{ij} D_{ij}$ .

### 3.2 Parameters of Interest and Threats to Their Identification

Equation (2) is a causal model that highlights our parameters of interest, as well as the threats to their identification in observational data. Of primary interest are the  $\tilde{\nu}_j$ 's, as these answer the causal question of how the outcome of the average student would change if she were to enroll in school  $j$  instead of school 0, i.e. the value-added of  $j$  over 0:  $\mathbb{E}_{\mathcal{I}}[Y_{ij} - Y_{i0}] = \mathbb{E}_{\mathcal{I}}[Y_{ij}] - \mathbb{E}_{\mathcal{I}}[Y_{i0}] = \nu_j - \nu_0 = \tilde{\nu}_j$ .

The threats to identifying  $\tilde{\nu}_j$  arise in the potential relationships between college enrollment choices  $D_{ij}$  and the two unobserved error terms in (2): the student fixed effect  $\alpha_i$ , and the realized match effect  $\tilde{\epsilon}_i$ . To clarify these two distinct threats, consider the simplest approach to identifying value-added: comparing the raw mean outcomes of students who attend different schools. This ‘‘raw outcomes’’ approach, widely used in popular college guides and in many states’ postsecondary funding formulas, can be implemented by regressing student outcomes  $Y_i$  on indicators for each college  $D_{ij}$ , omitting the reference school  $j = 0$ :

$$Y_i = b_0 + \sum_{j \neq 0} b_j D_{ij} + e_i, \tag{3}$$

where  $b_0 = \mathbb{E}[Y_i | D_{i0} = 1]$  and  $b_j = \mathbb{E}[Y_i | D_{ij} = 1] - \mathbb{E}[Y_i | D_{i0} = 1]$  by definition of linear regression. Plugging

in the potential outcome primitives from (1) yields the decomposition of interest for a given school  $j$ :

$$\begin{aligned}
b_j &= \mathbb{E}[Y_i|D_{ij} = 1] - \mathbb{E}[Y_i|D_{i0} = 1] \\
&= \underbrace{\tilde{\nu}_j}_{\text{Value-added}} + \underbrace{\mathbb{E}[\alpha_i|D_{ij} = 1] - \mathbb{E}[\alpha_i|D_{i0} = 1]}_{\text{Vertical selection bias}} + \underbrace{\mathbb{E}[\epsilon_{ij}|D_{ij} = 1] - \mathbb{E}[\epsilon_{i0}|D_{i0} = 1]}_{\text{Differential match bias}}. \tag{4}
\end{aligned}$$

Equation (4) shows that the regression parameter  $b_j$  in the raw outcomes approach can be biased away from the target causal parameter  $\tilde{\nu}_j$  by two different forces: 1) vertical selection bias, in which the types of students who sort into school  $j$  would have reaped higher (or lower) average outcomes than students at school 0 regardless of any causal impacts of the schools themselves, and 2) differential match bias, in which the average realized student-school match quality among students at school  $j$  may be higher (or lower) than the average realized match quality among students at school 0. We now address these two different threats to identification in turn.

### 3.3 Addressing Vertical Selection Bias: Proxying for Student Fixed Effects

The causal model in Equation (2) immediately suggests a student fixed effects approach to addressing the problem of vertical selection bias. Unfortunately, our setting does not feature a panel structure of outcomes driven by contemporaneous treatments that vary within person over time, like the earnings histories of individual workers who switch across firms or the test score histories of individual students who face a series of different teachers. Hence we cannot include indicators for each student to directly control for the fixed effects  $\alpha_i$ . Instead, we propose to proxy for  $\alpha_i$  at a coarser level with indicators for each distinct portfolio of college applications and admissions. Intuitively, the idea behind this approach is that when students choose a portfolio of applications, they provide a high-dimensional signal of their abilities, ambitions, and advantages, while admissions officers layer on an additional high-dimensional signal of their judgement about the student's potential through their collective acceptance decisions.

Concretely, consider the population projection of  $\alpha_i$  onto a mutually exclusive and exhaustive set of indicators for each admission portfolio  $A_{ip}$  indexed by  $p \in \mathcal{P}$ , omitting the base category  $p = 0$ :

$$\alpha_i = \phi_0 + \sum_{p \neq 0} \phi_p A_{ip} + \eta_i, \tag{5}$$

where  $\phi_0 = \mathbb{E}[\alpha_i|A_{i0} = 1]$ ,  $\phi_p = \mathbb{E}[\alpha_i|A_{ip} = 1] - \mathbb{E}[\alpha_i|A_{i0} = 1]$ , and  $\mathbb{E}[\eta_i|A_{ip}] = 0$  by construction. The key condition of the proxy approach is that a student's admissions portfolio  $A_{ip}$  must capture sufficient

variation in  $\alpha_i$  such that the residual within-portfolio variation,  $\eta_i$ , is uncorrelated with where students ultimately decide to enroll,  $D_{ij}$ . Specifically, we make the following mean-independence assumption:

**Assumption 1** *Admissions portfolios proxy for student fixed effects:*  $\mathbb{E}[\eta_i|D_{ij}, A_{ip}] = \mathbb{E}[\eta_i|A_{ip}] = 0$ .

Assumption 1 allows us to eliminate vertical selection bias by conditioning on observable admissions portfolios, which serve as coarse but sufficient proxies for unobservable student fixed effects:

$$\begin{aligned} & \underbrace{\mathbb{E}[\alpha_i|D_{ij} = 1, A_{ip}] - \mathbb{E}[\alpha_i|D_{i0} = 1, A_{ip}]}_{\text{Vertical selection bias within portfolio } p} \\ &= \mathbb{E}[\phi_0 + \phi_p + \eta_i|D_{ij} = 1, A_{ip}] - \mathbb{E}[\phi_0 + \phi_p + \eta_i|D_{i0} = 1, A_{ip}] \\ &= \mathbb{E}[\eta_i|D_{ij} = 1, A_{ip}] - \mathbb{E}[\eta_i|D_{i0} = 1, A_{ip}] \\ &= \mathbb{E}[\eta_i|A_{ip}] - \mathbb{E}[\eta_i|A_{ip}] = 0, \end{aligned}$$

where the first equation plugs in (5) and the third equation imposes Assumption 1.

We conduct a battery of empirical tests to probe the validity of Assumption 1, detailed in Section 4.2. We find remarkably little evidence of any remaining vertical selection bias after conditioning solely on students' admissions portfolios, both through balance checks and tests for omitted variable bias. These results suggest that students facing the same choice set of college admissions do not further sort into institutions systematically by vertical traits like ability, ambition, and advantage, lending empirical credence to the notion in Assumption 1 that admissions portfolios sufficiently proxy for student fixed effects and thus adequately address the threat of vertical selection bias in identifying college value-added.

### 3.4 Addressing Differential Match Bias: Canon, Theory, and Evidence

The canonical value-added approach, whether applied to teacher effects (Chetty et al., 2014a), K-12 school quality (Angrist et al., 2017), industry wage differentials (Krueger and Summers, 1988), firm wage premia (Abowd et al., 1999), regional health care supply (Finkelstein et al., 2016), or intergenerational mobility across neighborhoods (Chetty and Hendren, 2018), tends to focus on vertical selection bias as the central threat to identification. This literature usually rules out bias from horizontal match effects, the other potential threat revealed by the decomposition in (4), by eliminating match effects  $\epsilon_{ij}$  from the model entirely (e.g. Chetty et al., 2014a; Angrist et al., 2017), or by granting  $\epsilon_{ij}$  a simple existence as orthogonal shocks that are unknown, or at least not acted upon, at the time of sorting into treatments.

Our baseline specification below follows this value-added canon, though it is interesting to note that our

decomposition in (4) implies we can address differential match bias with a somewhat weaker assumption than is typically invoked. Instead of eliminating match effects entirely ( $\epsilon_{ij} \equiv 0$ ) or assuming students are completely oblivious to them when making enrollment choices ( $\{\epsilon_{ij}\}_{j \in \mathcal{J}} \perp \{D_{ij}\}_{j \in \mathcal{J}}$ ), we can allow for a particular form of student sorting that preserves mean-independence of the realized match effect  $\tilde{\epsilon}_i \equiv \sum_j \epsilon_{ij} D_{ij}$  across chosen schools  $D_{ij}$ , conditional on the admission portfolio  $A_{ip}$ :

**Assumption 2** *No differential match bias within admissions portfolios:*  $\mathbb{E}[\tilde{\epsilon}_i | D_{ij}, A_{ip}] = \mathbb{E}[\tilde{\epsilon}_i | A_{ip}] \equiv \bar{\epsilon}_p$ .

Note that differential match bias is the difference in average *realized* match quality among students who end up at school  $j$  vs. students who end up at school 0:  $\mathbb{E}[\epsilon_{ij} | D_{ij} = 1, A_{ip}] - \mathbb{E}[\epsilon_{i0} | D_{i0} = 1, A_{ip}]$ . Assuming that this difference is zero via Assumption 2 does not rule out systematic self-selection of students into schools at which they have higher match quality, since Assumption 2 does not impose  $\mathbb{E}[\tilde{\epsilon}_i | A_{ip}] = 0$ ; instead, it rules out *differential* magnitudes of such selection, on average, across students with the same portfolio who end up at different schools. That is, with UT-Austin as the reference school ( $j = 0$ ), students who choose to enroll at Texas Tech can do so because they tend to have positive match quality at Texas Tech ( $\epsilon_{ij} > 0$ ), and students who choose to enroll at UT-Austin can do so because they tend to have positive match quality at UT-Austin ( $\epsilon_{i0} > 0$ ). As long as the strength of this self-selection into different colleges averages up to the same magnitude  $\bar{\epsilon}_p$  among students with the same portfolio  $p$ , it will difference out in within-portfolio comparisons of student outcomes across campuses. In short, we simply note that Assumption 2, which is implicitly invoked by the vast majority of the value-added literature, need not shut down sorting on idiosyncratic match effects altogether, though it does require a particular form of such sorting that is symmetric across chosen treatments conditional on controls.

More importantly, we relax Assumption 2 and directly investigate the presence of match effects between students and colleges in Section 8 by allowing value-added to vary flexibly across several dimensions of student heterogeneity, including those often hypothesized to be key drivers of (mis)match effects—race, cognitive skills, and family income—as well as gender and non-cognitive skills. We find little evidence of systematic heterogeneity in value-added across most student subgroups, and our value-added estimates for the average student are very similar regardless of whether we include a rich set of interactions between college treatments and student covariates (see Figure A.5) to account for potential match effects. These results lend empirical support to the notion in Assumption 2 that realized match effects are not systematically different across campuses, and thus do not inject differential match bias into our main value-added estimates.

### 3.5 Identification and Implementation

We now have all the ingredients necessary to derive our baseline regression specification. From the causal model in (2), first plug in the projection from (5) to get

$$Y_i = \tilde{\kappa} + \sum_{j \neq 0} \tilde{\nu}_j D_{ij} + \phi_0 + \sum_{p \neq 0} \phi_p A_{ip} + \eta_i + \tilde{\epsilon}_i,$$

then add and subtract the average realized match quality for each admissions portfolio  $\mathbb{E}[\tilde{\epsilon}_i | A_{ip}] \equiv \bar{\epsilon}_p$ ,

$$Y_i = \tilde{\kappa} + \sum_{j \neq 0} \tilde{\nu}_j D_{ij} + \phi_0 + \sum_{p \neq 0} \phi_p A_{ip} + \eta_i + \bar{\epsilon}_0 + \sum_{p \neq 0} (\bar{\epsilon}_p - \bar{\epsilon}_0) A_{ip} + (\tilde{\epsilon}_i - \sum_p \bar{\epsilon}_p A_{ip}),$$

and finally collect terms:

$$\begin{aligned} Y_i &= \underbrace{\tilde{\kappa} + \phi_0 + \bar{\epsilon}_0}_{\equiv \ddot{\kappa}} + \sum_{j \neq 0} \tilde{\nu}_j D_{ij} + \sum_{p \neq 0} \underbrace{(\phi_p + \bar{\epsilon}_p - \bar{\epsilon}_0)}_{\equiv \ddot{\phi}_p} A_{ip} + \eta_i + \underbrace{(\tilde{\epsilon}_i - \sum_p \bar{\epsilon}_p A_{ip})}_{\equiv \ddot{\epsilon}_i} \\ \Rightarrow Y_i &= \ddot{\kappa} + \sum_{j \neq 0} \tilde{\nu}_j D_{ij} + \sum_{p \neq 0} \ddot{\phi}_p A_{ip} + \ddot{\epsilon}_i. \end{aligned} \tag{6}$$

Under Assumptions 1 and 2, the composite regression error  $\ddot{\epsilon}_i$  in (6) satisfies

$$\mathbb{E}[\ddot{\epsilon}_i | D_{ij}, A_{ip}] = \underbrace{\mathbb{E}[\eta_i | D_{ij}, A_{ip}]}_{=0 \text{ by Assumption 1}} + \underbrace{\mathbb{E}[\tilde{\epsilon}_i | D_{ij}, A_{ip}]}_{=\bar{\epsilon}_p \text{ by Assumption 2}} - \underbrace{\mathbb{E}[\sum_p \bar{\epsilon}_p A_{ip} | D_{ij}, A_{ip}]}_{=\bar{\epsilon}_p \text{ by definition}} = 0.$$

Hence, our baseline specification follows (6) and regresses outcomes  $Y_i$  on the set of college treatment indicators  $D_{ij}$ , omitting UT-Austin as the  $j = 0$  reference category, plus fixed effects for each distinct admission portfolio  $A_{ip}$ , omitting the reference category  $p = 0$ . Under Assumptions 1 and 2, the coefficients on  $D_{ij}$  identify our parameters of interest  $\tilde{\nu}_j$ : the causal value-added of each college  $j$  relative to UT-Austin.

## 4 Value-Added Estimates and Validation Exercises

### 4.1 Baseline Estimates and Comparison to Other Approaches

Figure 2 presents our baseline value-added estimates on BA completion (top panel) and earnings (bottom panel), along with estimates from two other common approaches for comparison.<sup>24</sup> Each plot begins on

<sup>24</sup>Appendix Tables B.1 and B.2 report the corresponding numerical estimates.

the left with raw outcome means at each of our thirty Texas public universities. The schools are ordered from top to bottom by these raw outcome means, which are measured relative to UT-Austin (signified by the vertical line at zero). We see large disparities in raw student outcomes across campuses: 82% of students who begin college at UT-Austin complete a BA within 6 years, which is over quadruple the rate of Texas Southern students, and UT-Austin mean earnings are \$55,975 8-10 years out from college entry, roughly double those of Texas Southern.

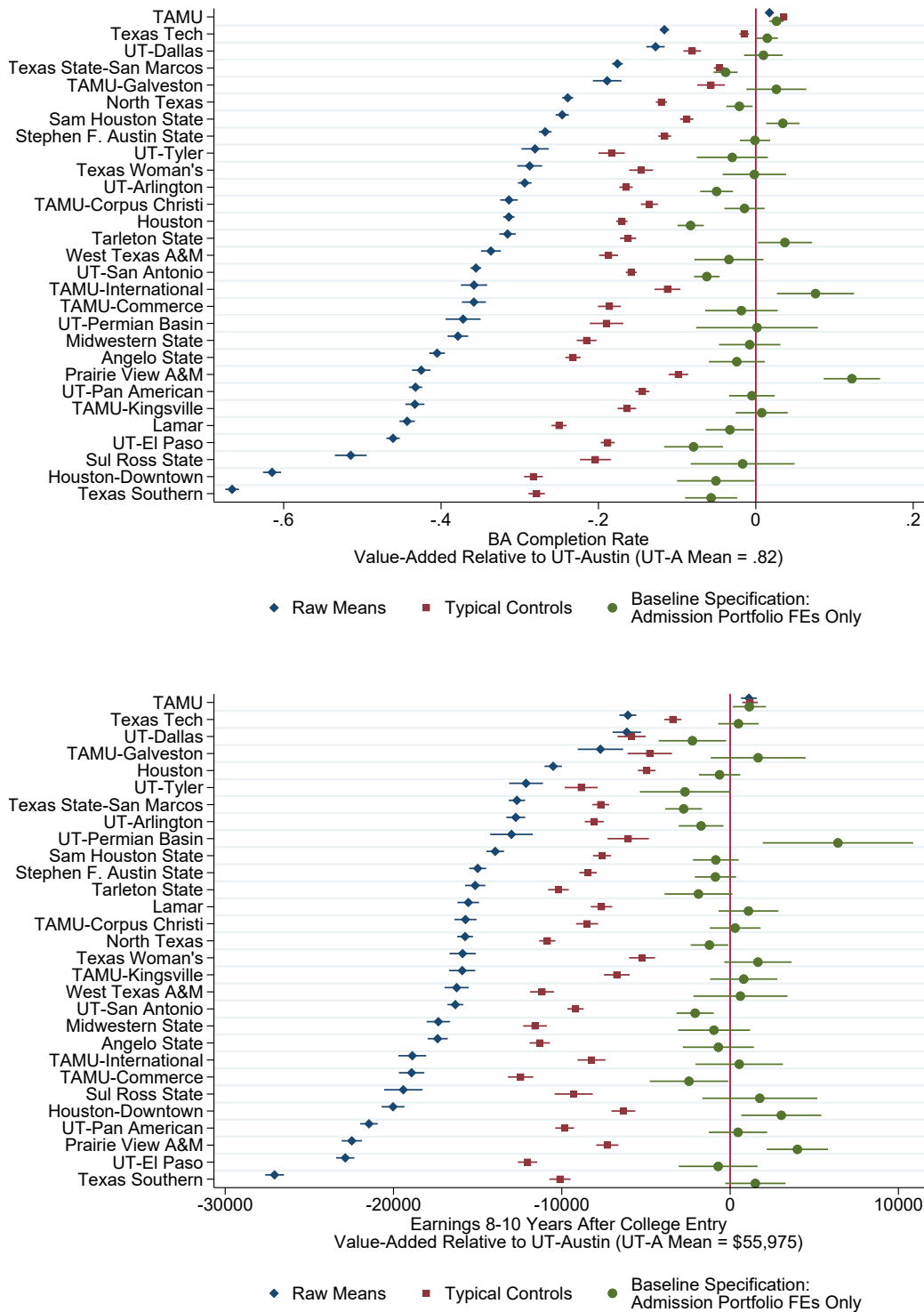
The estimates in the middle column of Figure 2 (squares) add a rich set of student covariate controls akin to much of the literature on the return to college quality: demographics (gender, race), family income (proxied by free/reduced price lunch status), high school academic preparation (10<sup>th</sup> grade test scores, advanced coursework, and an indicator for graduating in the top GPA decile), and behavioral measures of non-cognitive skills (high school attendance record, disciplinary infractions, and an indicator for ever being at-risk of dropping out). All of these covariates are highly predictive of both college choices and outcomes; controlling for them attenuates the mean outcome differences between most colleges and UT-Austin by roughly half, but interpreting the remaining differences as causal would still attribute substantial returns to attending the more selective colleges at the top of the distribution relative to those at the bottom.

Our baseline value-added estimates, which control solely for admission portfolio fixed effects as in Equation (6), appear farthest to the right of Figure 2 (in circles). Remarkably, they cluster around zero: students who face the same choice sets of college admissions tend to have relatively similar graduation and earnings outcomes regardless of where they end up enrolling. This suggests that the systematic selection of students into colleges, rather than causal college value-added within admission sets, is the dominant factor driving the raw outcome disparities we observe across colleges, both in terms of cardinal magnitudes and ordinal rankings. As we quantify in Sections 5 and 6 below, the causal value-added gradient is still meaningful, though it is substantially less steep than the raw outcomes suggest, and it does not correspond well with traditional orderings like selectivity.

## 4.2 Validating the Matched Applicant Approach

Our baseline value-added estimates in Figure 2 control solely for admission portfolio fixed effects. These estimates may be biased if students with the same portfolio of admissions further sort into institutions systematically by unobserved potential, which would violate Assumption 1. While we cannot definitively rule out additional sorting on unobservables, we can make use of our extensive set of observable student covariates, all of which strongly predict outcomes even within admissions portfolios, to check for balance

Figure 2: Baseline Value-Added Estimates and Comparison to Other Approaches



Notes: Each set of point estimates and robust 95% confidence intervals come from regressions of individual student outcomes on college treatment indicators, omitting UT-Austin as the reference treatment (signified by the vertical line at zero). The UT-Austin outcome mean appears in parentheses below each plot. All specifications control for cohort fixed effects. The *Raw Means* specification controls for nothing else. The *Typical Controls* specification adds controls for demographics (gender, race, FRPL), high school academic preparation (10th grade test scores, advanced coursework, and top high school GPA decile indicator), and behavioral measures of non-cognitive skills (high-school attendance, disciplinary infractions, and an indicator for ever being at risk of dropping out). The *Baseline Specification* controls solely for college admission portfolio fixed effects (and cohort fixed effects). See Appendix Tables B.1 and B.2 for the corresponding numerical estimates.

across treatments and probe the potential for omitted variable bias with successively richer control sets. We also explore whether alternate specifications of the admissions portfolio controls deliver similar estimates.

#### 4.2.1 Balance Across College Treatments

Figure 3 directly investigates whether students with the same admissions portfolio further sort into colleges systematically by ability. The figure presents estimates from specifications similar to those in Figure 2, except that “outcomes” in the regressions here are pre-college measures of student ability unused in college admissions: 10th grade standardized test scores in math (top panel) and reading (bottom panel). The leftmost estimates (diamonds) simply regress these scores on our college treatment dummies, recovering raw means across campuses relative to UT-Austin. Similar to Figure 2, these raw means vary dramatically across colleges, but controlling solely for admission portfolio fixed effects (circles) starkly reduces these ability differences to rough balance across college treatments. Here, balance implies that students with a given portfolio who enroll at UT-Austin tend to have very similar measured ability as students with the same portfolio who enroll elsewhere. This is consistent with Assumption 1 that residual variation in student potential within admission portfolios is uncorrelated with where students ultimately enroll.

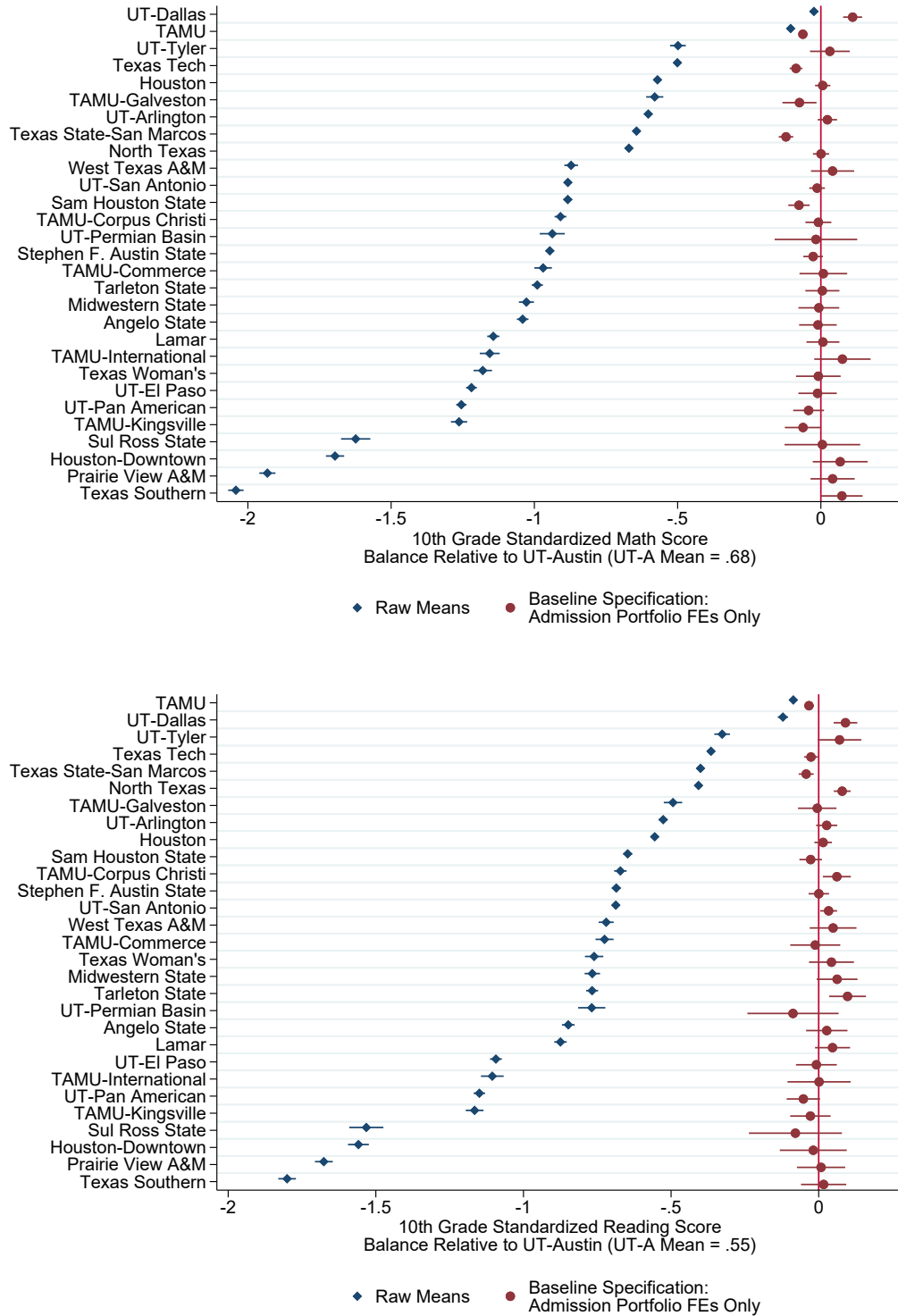
#### 4.2.2 Robustness to Omitted Variables

Admissions portfolio controls nearly eliminate observable differences in ability across our treatment schools, though Figure 3 shows that some minor deviations remain. Thus, to move from our baseline specification to our main specification, we add a core set of control variables, which are also the dimensions of student heterogeneity along which we investigate match effects in Section 8: categorical indicators for gender, race, and low family income (FRPL), and continuous measures of cognitive skills (10th grade test scores) and non-cognitive skills (high school attendance).

Figure 4 shows that adding these core controls barely moves the value-added estimates, suggesting that the minor deviations from perfect balance in Figure 3 do not translate into meaningful bias. The additional results in Figure 4 demonstrate robustness of our value-added estimates to further expansions of the control set. *Extra Control Set 1* includes additional pre-college measures of cognitive and non-cognitive skills: advanced high school coursework, disciplinary infractions, and an indicator for ever being at risk of dropping out. *Extra Control Set 2* further adds the indicator for graduating in the top decile of high school GPA, as well as fixed effects for each of the 1,457 different high schools from which our students graduate, flexibly accounting for the environmental influences of the schools, neighborhoods, and local

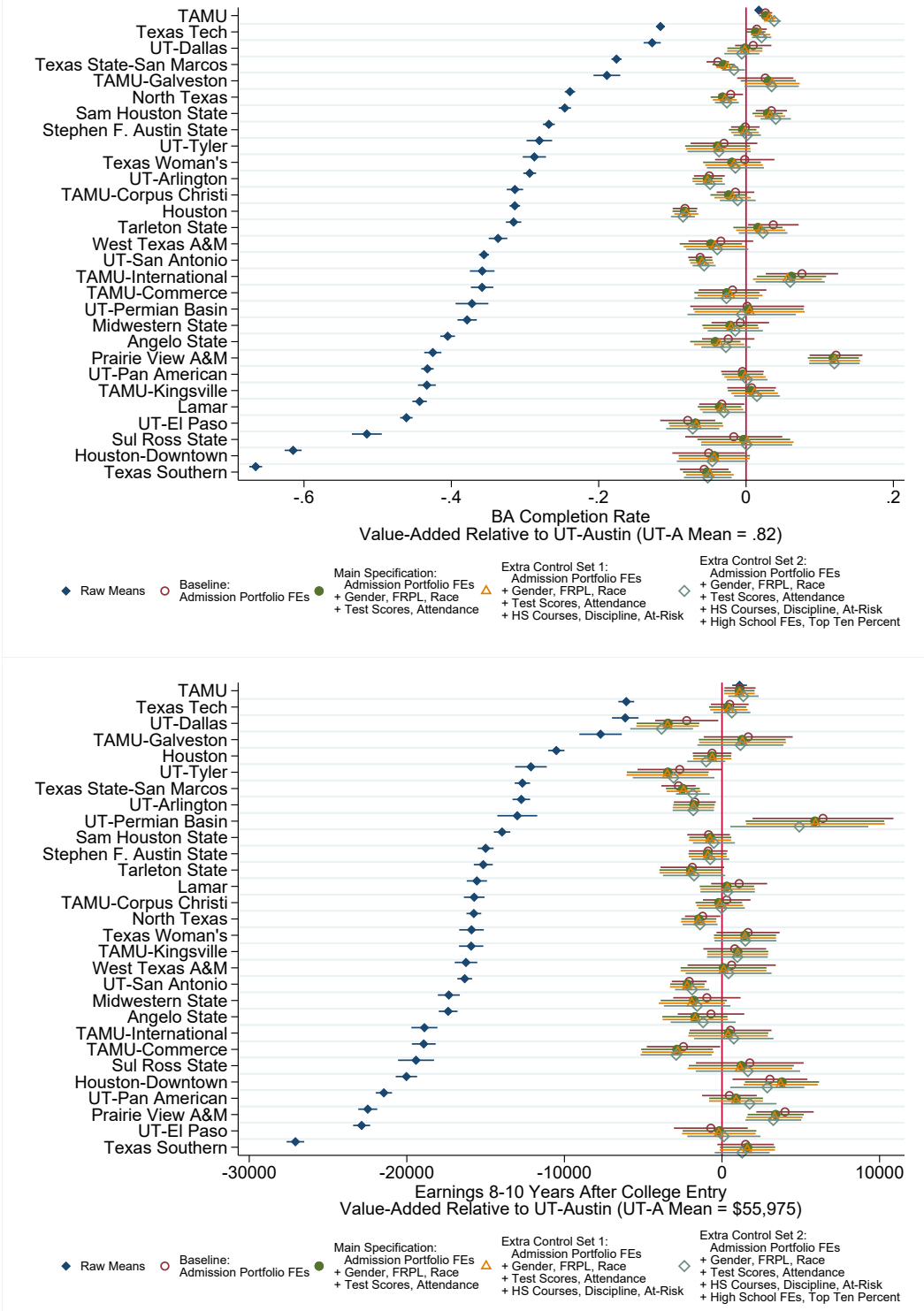


Figure 3: Validating the Matched Applicant Approach: Ability Balance across College Treatments



Notes: Each set of point estimates and robust 95% confidence intervals come from regressions of individual student standardized test scores on college treatment indicators, omitting UT-Austin as the reference treatment (signified by the vertical line at zero). The UT-Austin mean appears in parentheses below each plot. The *Raw Means* specification controls only for cohort fixed effects. The *Baseline Specification* controls solely for college admission portfolio fixed effects (and cohort fixed effects). See Appendix Table B.3 for the corresponding numerical estimates.

Figure 4: Validating the Matched Applicant Approach: Robustness to Omitted Variables



Notes: Each set of point estimates and robust 95% confidence intervals come from regressions of individual student outcomes on college treatment indicators, omitting UT-Austin as the reference treatment (signified by the vertical line at zero). The UT-Austin outcome mean appears in parentheses below each plot. All specifications control for cohort fixed effects. The *Raw Means* specification controls for nothing else. The *Baseline Specification* controls solely for college admission portfolio fixed effects (and cohort fixed effects). The *Main Specification* adds our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. *Extra Control Set 1* adds controls for advanced high school coursework, disciplinary infractions, and an indicator for ever being at risk of dropping out. *Extra Control Set 2* adds fixed effects for every high school and an indicator for being in the top decile of high school GPA. See Appendix Tables B.1 and B.2 for the corresponding numerical estimates.

labor markets in which our students are embedded prior to college entry. All of these additional controls are strong predictors of student outcomes, even within admissions portfolios: moving from the baseline specification to the specification with all extra controls increases the adjusted  $R^2$  proportionally by about 60% (from .145 to .235 for BA completion, and from .089 to .142 for earnings). Meanwhile, the value-added estimates barely change. While this exercise cannot completely rule out bias from remaining unobservable confounders, the stability of the value-added estimates in Figure 4 at least fails to invalidate our research design across the expansive swath of potential confounders that we do observe.

### 4.2.3 Insufficiency of Simpler Admissions Portfolio Specifications

A natural question, prompted by the use of exhaustive fixed effects for every distinct admissions portfolio, is whether more parsimonious specifications would deliver similar value-added estimates. Appendix Figure A.2 highlights perils along at least two dimensions of simplification. First, the estimates in hollow triangles come from specifications that control simply for 29 additive indicators for applying to each school and 29 additive indicators for being admitted to each school (omitting UT-Austin), which reduces the number of portfolio control coefficients from thousands to just 58. This more restrictive approach eliminates information from interactions between application and admission decisions, and the results indicate that it substantially under-controls for selection: the estimates lay roughly halfway between the raw outcome means and our main causal estimates.<sup>25</sup>

Second, the estimates in hollow circles in Figure A.2 come from specifications that include fixed effects for each distinct portfolio of *applications* only, ignoring any information about admissions decisions. These estimates get closer to our main results, but they still appear to systematically under-control for selection, leaving larger outcome gaps between the selective flagships and the rest of the schools. Admissions decisions thus appear to reveal additional information about student potential beyond application behavior. This suggests that research designs relying solely on application controls may not fully correct for selection, and thus may overestimate the causal value-added of attending more selective institutions.

### 4.2.4 Robustness to Richer Specifications of Admissions Portfolios

Figure A.3 probes robustness to more demanding specifications of the admissions portfolio controls. First, limiting the analysis sample to the subset of students who are admitted to at least two colleges delivers

---

<sup>25</sup>Another related reduction of the number of portfolio controls would be to group all students who apply to four or more schools into one portfolio (e.g. Cunha and Miller, 2014); we find that this approach also appears to significantly under-control for selection, given the large heterogeneity in student ability within that pooled category.

nearly identical results as the full sample, since our value-added parameters are identified from variation in enrollment choices among this subset. Students with only one admission necessarily have no scope for variation in enrollment choices and thus do not contribute to identifying value-added in this approach.

The next set of estimates in Figure A.3 add information about rejections (applications that do not result in admissions) in the portfolio control. Recall that our main specification uses portfolios based on simplified binary entries for each college: *do not get admitted* or *apply and get admitted*, where *do not get admitted* could be due either to not applying or applying and getting rejected. The estimates in triangles in Figure A.3 use portfolios with all three possible entries for each college: *do not apply*, *apply and get rejected*, or *apply and get admitted*. The resulting estimates are nearly identical to those from our main specification. Thus, colleges' rejection decisions, and students' choices to apply to colleges that ultimately reject them, do not provide additional selection correction beyond students' portfolios of admissions.

As a final check for robustness to richer specifications, the last two sets of estimates in Figure A.3 interact our main admissions portfolios with additional covariates. The first specification interacts them with the indicator for being in the top decile of high school GPA, to address the concern that a given portfolio may convey different information if the student is eligible for the Texas Top Ten Percent automatic admissions program; our value-added estimates are unaffected. The second specification interacts the admission portfolios with our core set of student covariates—gender, race, low-income (FRPL), high school test scores, and high school attendance—to address the concern that a given portfolio may convey different information for students with different demographic backgrounds and skills. Again, the value-added estimates are virtually unaffected, except perhaps for our smallest college (UT-Permian Basin) where they shift somewhat but not in a statistically distinguishable way relative to our main specification.

### 4.3 Additional Robustness Checks: Alternate Earnings Measures & Missing Earnings

Our last set of checks probe the robustness of our earnings results to alternative earnings definitions and missing earnings. The top two panels in Appendix Figure A.4 scatter our main value-added estimates on earnings levels against value-added estimates on log earnings (left) and earnings rank within our sample of college-goers (right). Both alternate earnings definitions produce value-added estimates with correlations above 0.90 with our main estimates. We also investigate whether our results are affected by the age at which we measure earnings (third plot of Figure A.4); our main estimates of value-added on earnings at ages 27-29 feature a correlation of 0.84 with less precise estimates from the subsample of cohorts for whom we can extend earnings ages to 30-32.

Regarding missing earnings, the fourth plot in Figure A.4 shows a near-perfect correlation (0.98) between our raw earnings means at each college in our sample with the raw earnings means for those same colleges calculated by Chetty et al. (2020) using nationwide tax records, suggesting that our Texas-workers-only sample accurately captures earnings patterns across schools.<sup>26</sup> We also compute the correlation between a school’s causal value-added on earnings and its causal value-added on appearing in the earnings sample (fifth plot in Figure A.4) and find that it is roughly zero. This helps further assuage concerns that our estimates are biased by systematic selection out of the earnings sample. As a final check of this, the last plot in Figure A.4 shows that our main value-added estimates on BA completion, which are not conditioned on appearing in the earnings sample, are essentially identical (correlation 0.99) to estimates from the subsample of students who do appear in the earnings sample.

## 5 Distributional Magnitudes of Value-Added Across Colleges

The results in the previous section presented visual evidence that causal value-added varies much less across colleges than raw outcome means, with the two measures being imperfectly correlated. We now quantify these distributional magnitudes, taking into account the extra variability present in our value-added estimates from finite-sample estimation error.

We begin by decomposing our value-added estimates  $\hat{v}_j$  for each college  $j$  (relative to UT-Austin) into true (signal) value-added  $v_j$  plus orthogonal estimation error (noise)  $e_j$ :  $\hat{v}_j = v_j + e_j$ , where  $\mathbb{E}[e_j|v_j] = 0$ . The variance of our estimates across colleges,  $\mathbb{V}[\hat{v}_j] \equiv \sigma_{\hat{v}}^2$ , is thus the sum of the signal variance  $\mathbb{V}[v_j] \equiv \sigma_v^2$  and the noise variance  $\mathbb{V}[e_j] \equiv \sigma_e^2$ . This means we can subtract the noise variance from the variance of our estimates to identify the underlying signal variance of true value-added, and we can take the square root of both sides to work in standard deviation units:

$$\sigma_v = \sqrt{\sigma_{\hat{v}}^2 - \sigma_e^2}. \tag{7}$$

We estimate  $\sigma_{\hat{v}}^2$  as the sample variance of the value-added estimates across colleges, and we estimate  $\sigma_e^2$  as the average of the squared standard errors of those value-added estimates, with both calculations weighted by enrollment to reflect variability in the value-added experienced by enrolled students.

Table 2 presents estimates of the components of Equation (7). Panel A quantifies the dispersion in raw outcome means across colleges, which is precisely measured and economically large: one standard

---

<sup>26</sup>Our mean earnings levels tend to be proportionally lower across all schools given that we measure earnings at slightly younger ages compared to most of the cohorts in Chetty et al. (2020).

Table 2: Distributional Magnitudes of College Value-Added, Accounting for Estimation Error

	BA Completion	Earnings
<i>Panel A: Raw Outcome Means</i>		
Standard deviation of estimates across colleges	.179	8,070
Standard deviation of signal component	.179	8,065
Standard deviation of noise component	.004	276
<i>Panel B: Causal Value-Added Estimates</i>		
Standard deviation of estimates across colleges	.039	1,530
Standard deviation of signal component	.037	1,332
Standard deviation of noise component	.012	753
<i>Panel C: Relationships between Raw Outcome Means and Value-Added</i>		
Signal SD of causal value-added ÷ signal SD of raw outcome means	.207	.165
Correlation of VA estimate with raw outcome mean (uncorrected for noise)	.471	.176
Correlation of signal VA with raw outcome mean (corrected for noise)	.495	.203
Regression of school's value-added estimate on its raw outcome mean (SE)	.103 (.036)	.033 (.035)

Notes: All calculations weight colleges by student enrollment.

deviation (SD) spans 17.9 percentage points of BA completion, and \$8,065 of annual earnings 8-10 years out from college entry. In contrast, Panel B shows substantially less dispersion in causal value-added, which is also rather precisely measured: one SD in the signal distribution of value-added spans 3.7 percentage points of BA completion and \$1,332 of annual earnings, or roughly a 3% increase relative to sample mean earnings of \$44,834.<sup>27</sup> As a gauge of the market-wide quality gradient, these numbers imply (under a normal distribution of value-added) a gain of roughly 12.3 percentage points of BA completion and \$4,440 of annual income by moving from the 10<sup>th</sup> to the 90<sup>th</sup> percentiles of the college distribution.

The signal SD of college value-added for each outcome is around one-fifth of the SD in raw outcome means, as calculated in the first row of Panel C. As for their joint distribution, the third row of Panel C shows signal correlations of .495 for BA completion and .203 for earnings, which reflect only modest corrections for attenuation bias from estimation error.<sup>28</sup> Finally, to get a sense of the magnitudes of these relationships in outcome units, the fourth row of Panel C presents coefficients from a bivariate regression of a school's value-added estimate on its raw outcome mean. For BA completion, a 10 percentage point increase in the raw graduation rate predicts an increase in causal value-added of just 1 percentage point;

<sup>27</sup>Since students complete an average of 2.89 years of college, a one SD increase in earnings value-added roughly boosts earnings by just over 1% per year enrolled. This is quantitatively similar to Chetty et al. (2014b)'s estimate of a 1.3% earnings boost of having a teacher with one SD better test score value-added for one year of elementary or middle school.

<sup>28</sup>We can correct for estimation error in the correlation between our value-added estimates  $\hat{v}_j$  and precisely measured institutional characteristics  $x_j$  (in this case school  $j$ 's raw outcome mean) by multiplying their raw correlation by  $\frac{\sigma_{\hat{v}}}{\sigma_v}$ , since

$$Corr(v, x) = \frac{Cov(v, x)}{\sigma_v \sigma_x} = \frac{Cov(\hat{v} - e, x)}{\sigma_v \sigma_x} = \frac{Cov(\hat{v}, x)}{\sigma_v \sigma_x} = Corr(\hat{v}, x) \frac{\sigma_{\hat{v}}}{\sigma_v}.$$

for earnings, a \$10,000 increase in the raw earnings mean predicts an increase in causal value-added of just \$330. All of the above results help quantify the visual intuition from the plots in Section 4: while economically meaningful differences in value-added do exist across colleges, comparisons of raw outcome means are not strongly informative about either the ordinal rankings or the cardinal magnitudes of these differences in causal effectiveness.

## 6 Institutional Predictors of Value-Added

One takeaway from the preceding results is that a college’s raw outcome mean is an imperfect predictor of its causal value-added within admission sets, especially for earnings. We now ask whether other prominent college characteristics—including selectivity, non-peer inputs like instructional spending and the faculty/student ratio, and intergenerational mobility statistics—are useful predictors of value-added.

### 6.1 Selectivity vs. Value-Added

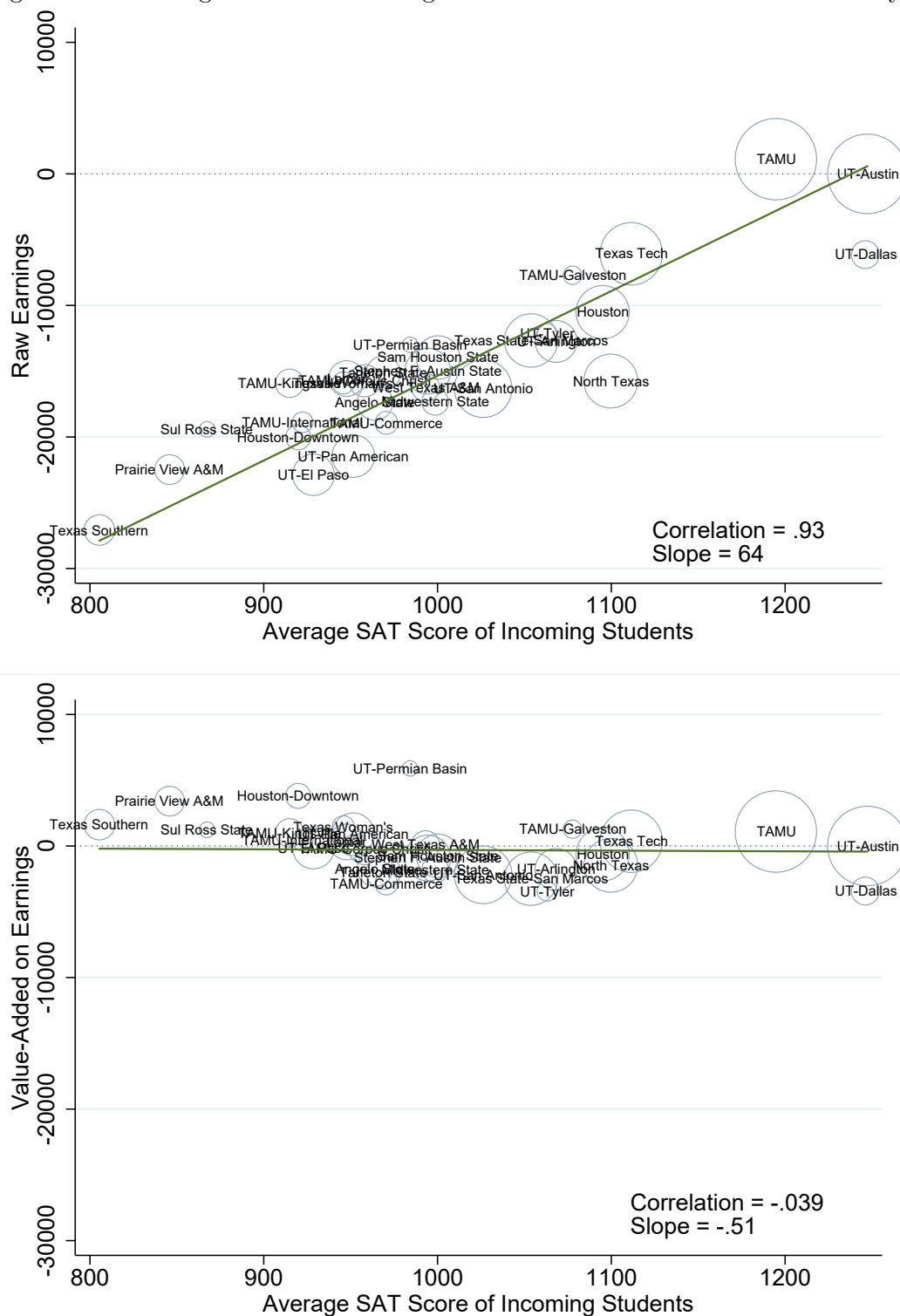
Selectivity, typically measured as the mean SAT score of a college’s incoming students, is a popular measure of vertical stratification in college guides, the popular press, and the academic literature on the returns to attending different types of colleges. The rich variation in selectivity across our sample colleges (recall Figure 1) provides a fruitful setting for exploring the relationship between selectivity and causal value-added in higher education. To set the stage, the top panel of Figure 5 shows that a college’s mean incoming SAT score is a very strong predictor of the *raw* mean earnings of its students, with a correlation of 0.93 and a regression slope of \$6,400 in annual earnings for each 100-point increase. In stark contrast, the bottom panel of Figure 5 shows that this relationship disappears with respect to value-added: selectivity is uninformative about a college’s causal impacts on admitted students’ earnings.<sup>29</sup>

Figure 6 presents further analyses from student-level regressions of earnings on the mean SAT score of the student’s college, akin to canonical specifications in the literature on the returns to college selectivity. The figure plots the regression coefficient on selectivity across several specifications. The top two omit our admission portfolio controls and suggest large returns to college selectivity. The remaining specifications control for admissions portfolio fixed effects and yield selectivity effects close to (and statistically indistinguishable from) zero. The final specification in Figure 6 is a robustness check: the zero return to selectivity still holds even within the subset of students who were admitted to UT-Austin, indicating that

---

<sup>29</sup>Correcting for estimation error in value-added has little impact on this correlation, consistent with the results in Table 2. Replacing each college’s mean SAT with its rejection rate, as another measure of selectivity, delivers nearly identical results.

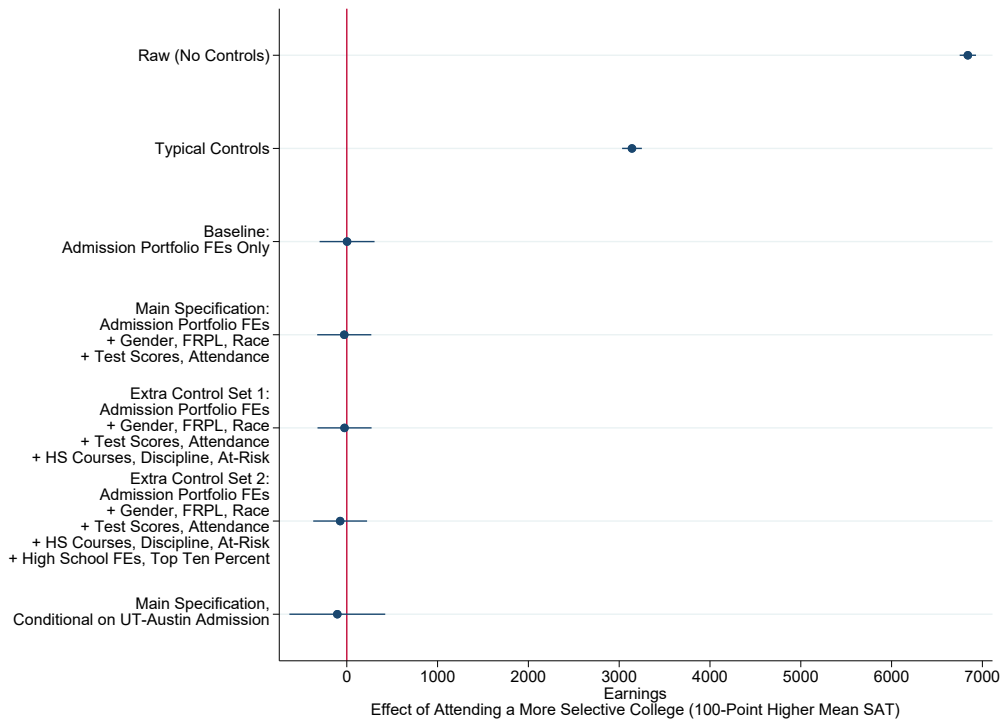
Figure 5: Predicting Raw Mean Earnings vs. Causal Value-Added with Selectivity



Notes: The top panel plots raw mean earnings at each college, relative to UT-Austin, against the average SAT score of incoming students at each college. The bottom panel replaces the vertical axis with our main value-added estimates from Section 4. Correlations, regression slopes, and circle sizes are weighted by student enrollment.



Figure 6: Student-Level Regression Estimates of the Return to Attending a More Selective College



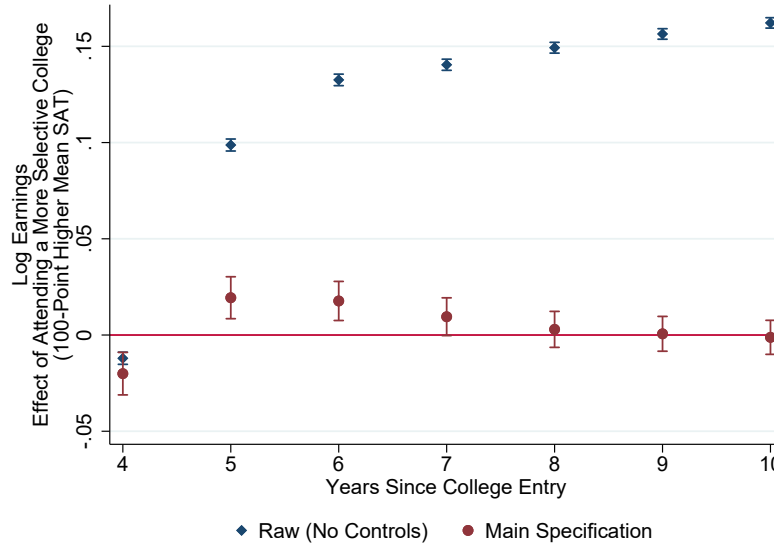
Notes: Each point estimate and robust 95% confidence interval comes from a regression of individual student outcomes on the mean incoming SAT score of the student’s college. The coefficients are scaled to correspond to a 100-point increase in mean SAT scores. All specifications control for cohort fixed effects. The *Raw Specification* controls for nothing else. The *Typical Controls* specification adds controls for demographics (gender, race, FRPL), high school academic preparation (10th grade test scores, advanced coursework, and top high school GPA decile indicator), and behavioral measures of non-cognitive skills (high school attendance, disciplinary infractions, and an indicator for ever being at risk of dropping out). The *Baseline Specification* controls solely for college admission portfolio fixed effects (and cohort fixed effects). The *Main Specification* adds our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. *Extra Control Set 1* adds controls for advanced high school coursework, disciplinary infractions, and an indicator for ever being at risk of dropping out. *Extra Control Set 2* adds fixed effects for every high school and an indicator for being in the top decile of high school GPA. The final estimate comes from running the Main Specification on the subsample of students who receive admission to UT-Austin. See Appendix Table B.6 for the corresponding numerical estimates.

the previous results are not driven by students choosing solely among less-selective institutions.

Interestingly, when plotting the early-career dynamics of the selectivity effect in Figure 7, we find that there does exist a causal earnings premium of about 2% per 100 SAT points in years 5-6 since college entry, but this premium fades out quickly to generate the stable null return in Figure 6 measured in years 8-10. This fleeting selectivity premium could be generated by a modest employer learning channel (e.g. Farber and Gibbons, 1996; Altonji and Pierret, 2001) or other early labor market frictions that give students from more selective colleges an initial advantage (e.g., campus recruiting, alumni networks) which fades quickly as students with similar potential from less selective colleges catch up.

These null results on selectivity, in tandem with the results in Section 5 and below in Section 6.2, help to resolve a long-standing controversy in the literature on whether meaningful economic returns accrue to

Figure 7: Early Career Dynamics of the Return to College Selectivity



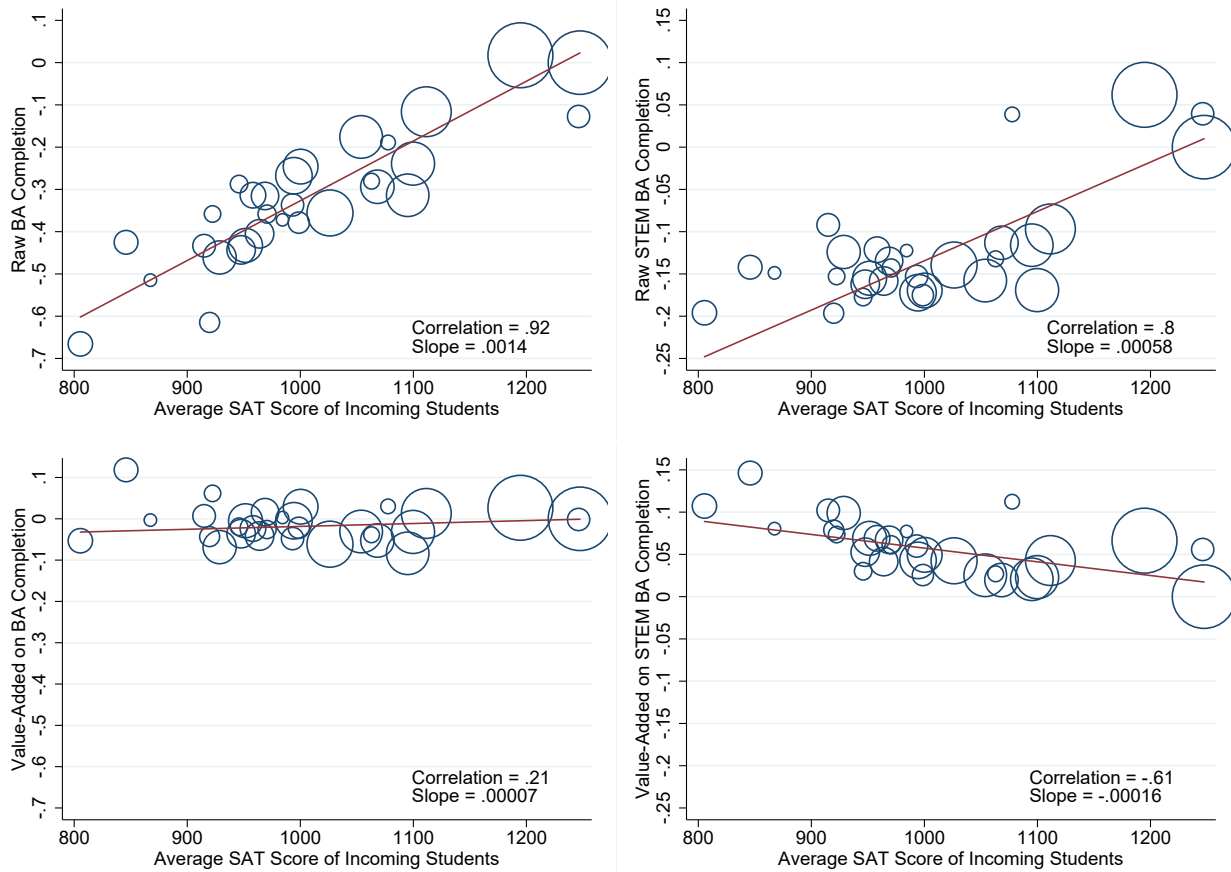
Notes: Each point estimate and robust 95% confidence interval comes from a separate regression of individual log student earnings, measured at a given number of years since college entry, on the mean incoming SAT score of the student’s college. The coefficients are scaled to correspond to a 100-point increase in mean SAT scores. All specifications control for cohort fixed effects. The *Raw Specification* controls for nothing else. The *Main Specification* controls for college admission portfolio fixed effects and our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance.

students who attend higher “quality” colleges. A majority of papers in this vein—e.g., Brewer et al. (1999), Black and Smith (2006), Long (2008), and Dillon and Smith (2018)—have found consistent evidence for strong returns, while Dale and Krueger (2002, 2014) stand out as a notable exception. Our results highlight the importance of carefully defining college “quality,” and provide some insight on why different strands of the literature have reached different conclusions. While we confirm Dale and Krueger (2002, 2014)’s main conclusion that college selectivity *per se* yields no meaningful economic returns (beyond a short-lived initial premium), this does not imply that college-level value-added differentials—a more flexible notion of quality—are absent. Rather, we document variation in causal value-added across colleges, and simply show that selectivity is orthogonal to it. More broadly, by letting each college have its own unique impact on student outcomes, our value-added approach lets the data determine the ordering of the quality space, rather than imposing an ex-ante ordering based on a single-dimensional college observable.<sup>30</sup>

To conclude our results on selectivity, Figure 8 explores two additional outcomes: completing any BA, and completing a BA in a STEM field. The two left panels show that, similar to earnings, raw BA completion rates exhibit a very strong correlation with selectivity (top left), but this relationship weakens dramatically when replacing raw outcomes with causal value-added (bottom left). A 100-point increase

<sup>30</sup>Black and Smith (2006) emphasize the need for multiple proxies to mitigate measurement error in univariate constructions of college quality. In the following subsection we show that our causal value-added estimates do covary with non-peer college inputs, while at the same time, they do not correspond perfectly to any one-dimensional observable college covariates.

Figure 8: Predicting Raw Degree Outcomes vs. Causal Value-Added with Selectivity



Notes: The top two panels plot raw BA completion rates (left) and raw STEM BA completion rates (right), relative to UT-Austin, against the average incoming SAT score at each college. STEM BA completion is not conditioned on completing a BA. The bottom two panels replace the vertical axes with value-added estimates from our main specification, which regresses individual student outcomes on college treatment indicators (with UT-Austin as the reference treatment at zero), controlling for admission portfolio fixed effects and our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. Correlations, slopes, and circles are weighted by student enrollment.

in incoming SAT scores predicts a 14 percentage point increase in raw BA completion, but less than 1 percentage point in value-added. The two right panels of Figure 8 reveal an even starker reversal for STEM degree completion. While raw STEM BA completion rates (top right) are also strongly positively correlated with selectivity (driven especially by the large selective flagships), the bottom right panel shows that this correlation does not simply attenuate for value-added, but actually strongly reverses in sign to a negative correlation of -0.61. STEM value-added increases roughly linearly at a rate of 1.6 percentage points for every 100-SAT-point *decline* away from UT-Austin, the most selective college in our sample.

## 6.2 Non-Peer Inputs vs. Value-Added

The weak correlation between selectivity and value-added suggests that peer quality is not an essential input in causal college effectiveness. We now explore whether non-peer college inputs, like instructional spending

Table 3: Correlations of Non-Peer College Inputs with Value-Added

	BA Completion	Earnings
<i>Non-Peer College Inputs: Correlation with Causal Value-Added</i>		
Instructional expenditures per student	.342	.317
Academic support expenditures per student	.158	.288
Student services expenditures per student	.295	.076
Share of faculty who are full-time	.371	.450
Share of faculty who are tenured or on tenure-track	.267	.411
Average faculty salary	.082	.090
Faculty/student ratio	.433	.433
Share of degrees in STEM fields	.332	.422

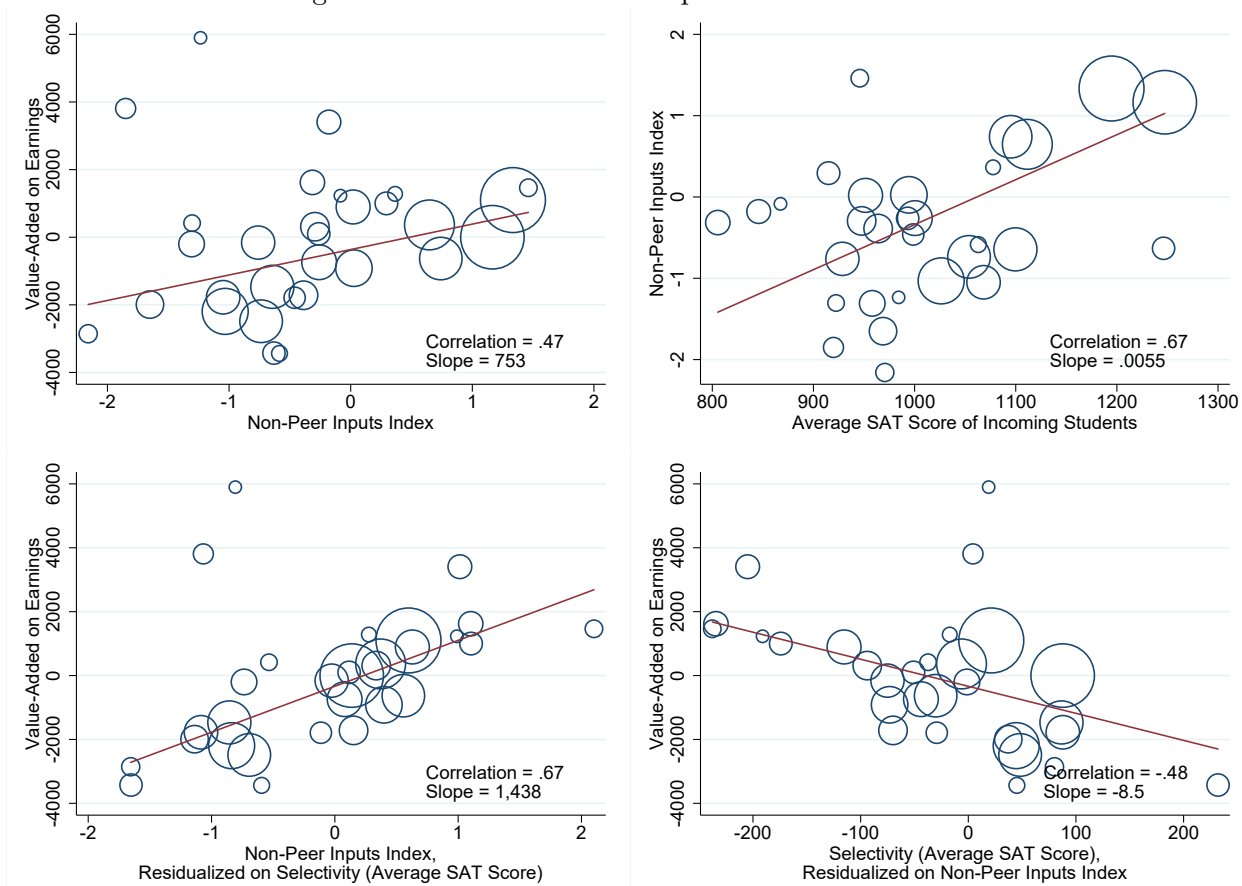
Notes: This table presents correlations between college-level measures of non-peer inputs and our main causal value-added estimates from Section 4. Non-peer input measures come from IPEDS and are averaged across available years from 2000 to 2017. Correlations are weighted by student enrollment.

and faculty characteristics, manage to covary positively with value-added. Table 3 presents correlations to this end. Among core postsecondary educational expenses tallied by IPEDS, instructional expenditures per student modestly correlate with value-added on both BA completion (.342) and earnings (.317), while expenditures on academic support and student services offer a mixture of somewhat weaker correlations. Among faculty characteristics, we tend to see stronger predictions for earnings value-added: the share of faculty who are full-time, the share of faculty who are tenured or on the tenure-track, and the faculty-to-student ratio feature healthy correlations with earnings value-added in the range of .41-.45. Average faculty salary, on the other hand, is not a strong predictor of value-added on either outcome. Finally, as a crude measure of differences in curricular inputs across campuses, we see that the share of degrees granted in STEM fields also covaries positively with value-added, especially on earnings, a pattern we will revisit in Section 7 when studying relationships between value-added on STEM degrees and earnings.

Since many of the input measures in Table 3 are correlated with each other, as well as with selectivity, we explore conditional relationships of peer and non-peer inputs with value-added in Figure 9. The top left panel plots earnings value-added against a single index of non-peer inputs, constructed as the predicted factor from an enrollment-weighted one-factor model of instructional expenditures, full-time faculty share, tenure-track faculty share, and faculty-student ratio. Reflecting a distillation and tightening of the component correlations from Table 3, we see an overall correlation of 0.47 between earnings value-added and this non-peer inputs index, and the slope implies that a one standard deviation increase in non-peer inputs across colleges predicts \$753 in additional earnings value-added.

The top right panel of Figure 9 shows that peer and non-peer inputs are positively correlated, but not perfectly so: there are plenty of off-diagonal colleges investing more or less in non-peer inputs than their mean student SAT score would predict, so the bottom panels of Figure 9 exploit this conditional

Figure 9: Peer vs. Non-Peer Inputs and Value-Added



Notes: The top left panel plots earnings value-added against an index of non-peer inputs, constructed as the predicted factor from an enrollment-weighted one-factor model of instructional expenditures, full-time faculty share, tenure-track faculty share, and faculty-student ratio. The top right panel plots this non-peer inputs index against the average incoming SAT score at each college. The bottom left panel plots earnings value-added against the residuals of a college-level regression of the non-peer inputs index on average SAT scores. Likewise, the bottom right panel plots earnings value-added against the residuals of a college-level regression of average SAT scores on the non-peer inputs index. All correlations, regressions, and circles are weighted by student enrollment.

variation. The bottom left plot shows that the positive relationship between value-added and non-peer inputs strengthens appreciably when controlling for selectivity: the partial correlation jumps to 0.67, and the slope nearly doubles to \$1,438 in extra predicted value-added for each standard deviation increase in non-peer inputs, residualized on selectivity. The bottom right panel shows the joint implication of this result and those from the previous subsection: selectivity, residualized on non-peer inputs, is actually a negative predictor of earnings value-added, perhaps reflecting a within-campus competition channel that becomes more apparent when comparing colleges with different peer composition but similar non-peer resources. These results further challenge popular notions of measuring college “quality” as positively-weighted indices of peer and non-peer inputs: although peer and non-peer inputs broadly move together, they each offer contrasting partial correlations with causal college value-added.

### 6.3 Intergenerational Mobility Statistics vs. Value-Added

Finally, we explore the potential for college-level intergenerational mobility statistics to serve as predictors of causal value-added. College-level mobility measures—the likelihood that students from disadvantaged backgrounds at a given college end up with relatively better outcomes as adults—have gained recent prominence in debates over whether higher education mitigates or exacerbates economic inequality. A key question regarding these measures is to what extent they reflect the causal effectiveness of particular colleges at boosting outcomes of disadvantaged students, versus the ability of such colleges to simply select disadvantaged students with high pre-existing potential. Chetty et al. (2020) estimate a comprehensive set of college mobility statistics using nationwide administrative tax records linking parental income, child income, and child college attendance. They measure mobility at college  $j$  as the probability that a child who attends  $j$  makes it into the top 20% of the national earnings distribution, conditional on that child having parents in the bottom 20% of the earnings distribution.<sup>31</sup>

Our data unfortunately do not include a continuous measure of parental income, nor do they include nationwide earnings, so we cannot reproduce Chetty et al. (2020)’s exact upward mobility measures with our microdata. The top panel of Figure 10, however, demonstrates that we can get rather close: instead of conditioning on students from the bottom family income quintile, we can condition on eligibility for free or reduced price lunch (FRPL) in high school, and instead of measuring the probability of making it into the top quintile of national earnings, we can measure the probability of making it into the top tercile of earnings among our Texas college-goers. This analogue measure in our data has a correlation of 0.96 with Chetty et al. (2020)’s measure from national tax data, suggesting that both samples are capturing the same patterns in raw intergenerational mobility across colleges.<sup>32</sup>

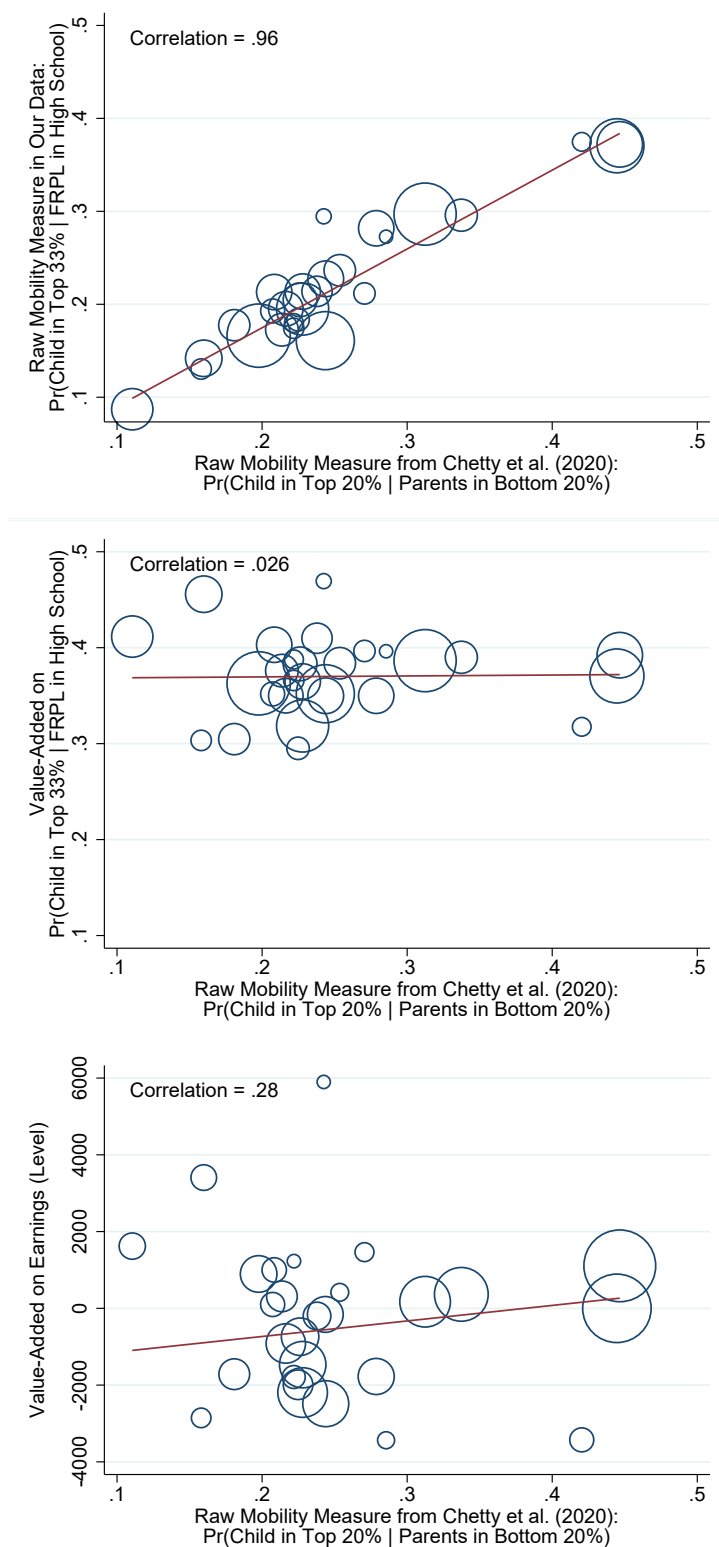
Having established common measurement of the raw mobility patterns, the middle panel of Figure 10 investigates to what degree a college’s raw mobility measure is predictive of its causal ability to transform disadvantaged students into high earners. To do this, we first estimate each college’s causal value-added on our analogue mobility measure by running our main value-added specification on the subsample of low-income (FRPL) students, and specifying the outcome as an indicator for whether the student makes it into the top tercile of income earners. We then plot these causal mobility estimates against the Chetty

---

<sup>31</sup>Chetty et al. (2020) denote this term the “success rate,” which they then multiply by an “access” measure (the fraction of students at each college who actually come from families in the bottom quintile) to generate their “mobility rate” metric. We focus on the success rate, since this is the component of the mobility rate that captures variation in the outcomes of observably similar students across colleges and is thus readily comparable to value-added.

<sup>32</sup>Chetty et al. (2020) combine some multi-campus university systems into single observational units, which in Texas has the consequence of merging TAMU-Galveston with TAMU-College Station, as well as Houston-Downtown with the larger University of Houston campus. For comparability in this subsection we combine the two schools in each pair as an enrollment-weighted average of their individual estimates.

Figure 10: Intergenerational Mobility Statistics vs. Value-Added



Notes: The horizontal axis in each plot is the college-level intergenerational mobility statistic reported by Chetty et al. (2020): the probability that a student at a given college reaches the top quintile of earnings conditional on having parents in the bottom quintile of earnings, estimated using national tax records. The vertical axis in the top panel is the closest analogue to this measure in our Texas sample: the raw probability that a student reaches the top income third of Texas college-goers conditional on being eligible for free or reduced price lunch in high school. The vertical axis in the middle panel swaps out the raw mobility measure for its causal analogue: a college's value-added on that probability, estimated using our main specification but limited to the subsample of FRPL students. The bottom panel replaces the vertical axis with our main earnings value-added measure from Section 4. Correlations and circles are weighted by student enrollment.

et al. (2020) raw mobility measures in the middle panel of Figure 10.<sup>33</sup> The plot reveals a correlation of roughly zero: a college’s raw intergenerational mobility statistic is an uninformative predictor of its causal value-added in lifting disadvantaged students to the top of the income distribution. This particular outcome measure may miss causal impacts across less ambitious leaps in the income distribution, however, so in the bottom panel of Figure 10 we swap in our main measure of value-added on average earnings levels, which exhibits a modest positive correlation of 0.28 with Chetty et al. (2020)’s raw mobility statistic.<sup>34</sup>

The lack of a strong relationship between raw mobility statistics and causal earnings effects across colleges is consistent with the results from Figure 2: the “typical controls” approach to estimating college value-added, which includes controls for family income akin to the raw mobility statistics, appears to substantially under-correct for selection, yielding results that differ substantially from our causal estimates derived from the matched applicant approach. Because the raw intergenerational mobility statistics control solely for family income, they are at risk of reflecting other systematic student differences across campuses, like levels of academic preparation and ambition, rather than causal college impacts.<sup>35</sup>

## 7 Potential Mechanisms: Relationships Across Student Outcomes

The preceding sections have focused on earnings and degree completion as separate outcomes on which colleges may add value. We now explore the interrelationships of value-added on these and other key outcomes to illustrate some of the potential mechanisms through which colleges shape the path from enrollment to earnings. To address attenuation bias in these relationships from measurement error in value-added, we construct Empirical Bayes (EB) forecasts (e.g., Robbins, 1956) that shrink each college’s value-added estimate towards the mean in proportion to its imprecision.<sup>36</sup> This issue turns out to be of minimal quantitative importance in our case, but we use EB-shrunk regressors to mitigate concerns about measurement error problems that generally affect the types of analyses we conduct below. See Appendix A.1 for further details on how we construct the EB forecasts.

---

<sup>33</sup>For this exercise, we add back in the mean outcome value at UT-Austin (the omitted school) to make the levels comparable to the Chetty et al. (2020) raw mobility measures on the horizontal axis.

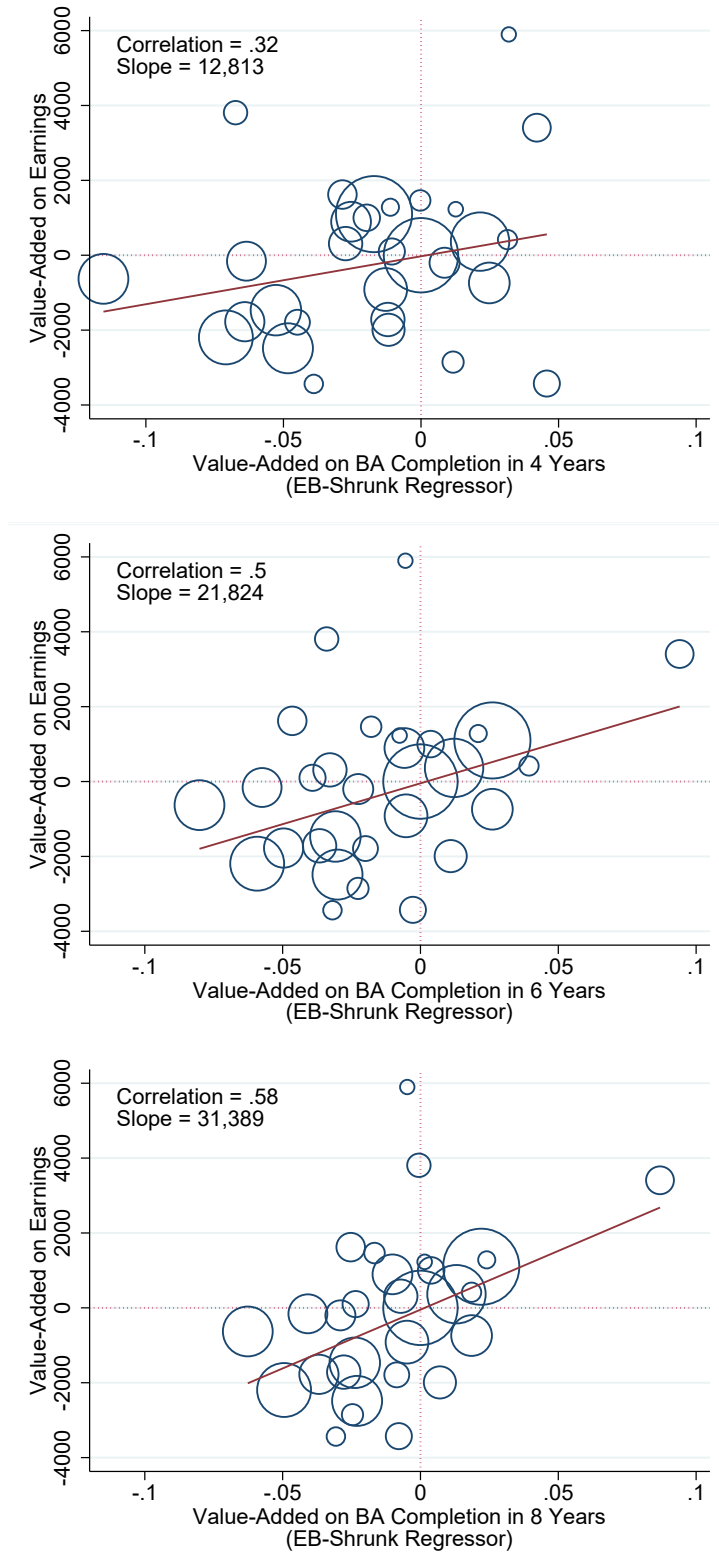
<sup>34</sup>This main measure of value-added on earnings levels is estimated on our full sample of students; using value-added estimates from the subsample of low-income (FRPL) students yields similar results, given our findings below in Section 8 that value-added does not appear to vary systematically by low-income status.

<sup>35</sup>Chetty et al. (2020) also estimate college mobility statistics that control for student SAT/ACT scores (in addition to family income) but are unable to publicly disclose these measures for comparison at the level of individual colleges. Chetty et al. (2020) further implement a specification in the spirit of Dale and Krueger (2002, 2014) that controls for SAT/ACT score sending behavior as a proxy for college applications, which they find only modestly attenuates outcome differences across colleges, contrasting with our results as well as those of Dale and Krueger (2002, 2014).

<sup>36</sup>Appendix Figure A.1 plots these EB forecasts against our value-added estimates. Since the value-added estimates for most colleges are relatively precise and/or already close to the mean, their shrunk EB forecasts are rather similar.



Figure 11: Value-Added on Earnings vs. Value-Added on BA Completion in 4, 6, 8 Years



Notes: The vertical axis of each plot is our main earnings value-added estimate from Section 4. The horizontal axes are the Empirical Bayes shrunk forecasts (see Appendix A.1) of value-added on BA completion within a given number of years, which correct the regression slopes for mild attenuation bias from estimation error in the regressor. Correlations, regression slopes, and circles are weighted by student enrollment.

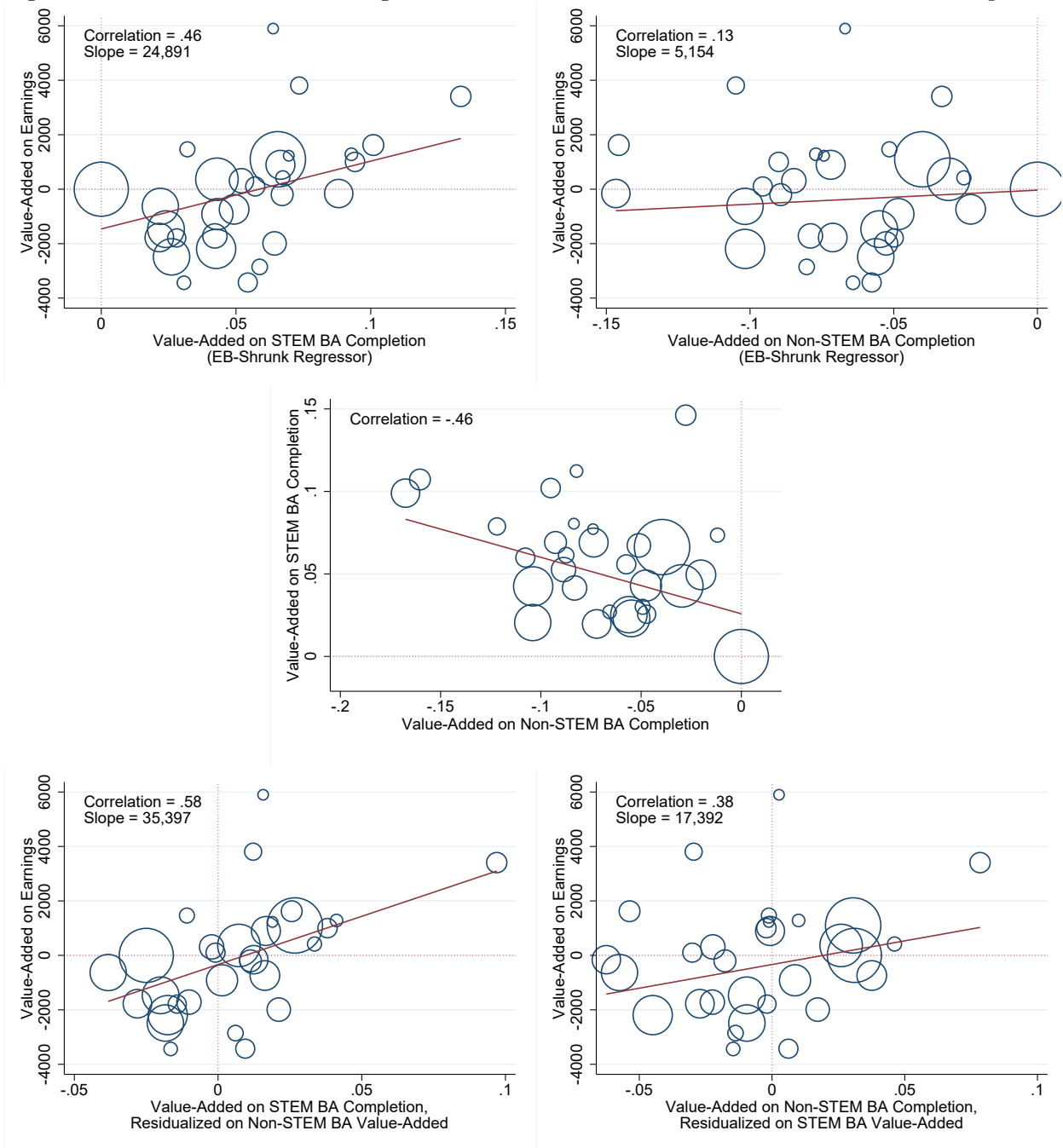
## 7.1 Value-Added on BA Completion vs. Value-Added on Earnings

We start in Figure 11 with the relationship between a college's value-added on BA completion (horizontal axis) and its value-added on earnings (vertical axis). The top panel shows a moderate positive correlation of 0.32 when measuring BA completion according to the "on-time" criterion of finishing within 4 years of entry. The middle panel extends the completion window to 6 years, and the bottom panel to 8 years. Each extension both tightens the distribution of BA value-added and strengthens its relationship with earnings value-added, with the slopes increasing from \$1,281 to \$2,182 to \$3,139 in additional earnings value-added associated with a 10 percentage point increase in BA value-added. Colleges are therefore more dispersed in their effectiveness at boosting on-time completion relative to eventual completion, and strict measures of on-time graduation may underestimate a college's ultimate value-added in the labor market.

## 7.2 Unpacking BA Completion: STEM vs. Non-STEM Degrees

We next unpack BA completion by STEM versus non-STEM degrees to explore whether value-added on completing different majors has different predictions for earnings effects. The top two panels of Figure 12 shows that a college's value-added on completing a STEM degree (left) has a strong correlation with its earnings value-added, while value-added on non-STEM completion (right), in contrast, has almost no bivariate relationship with earnings value-added. The middle panel shows that STEM and non-STEM value-added are negatively correlated, however, suggesting that colleges may face tradeoffs across fields in boosting degree completion. Exemplifying this to the extreme, UT-Austin at the (0,0) origin has the lowest value-added on STEM completion, as we also saw in Section 6, but simultaneously has the *highest* value-added on non-STEM degree completion. Simple bivariate correlations like the top right panel of Figure 12 may thus underestimate the earnings gains of producing more non-STEM degrees, since this is correlated with fewer STEM degrees in the cross-section. The bottom right panel of Figure 12, consistent with this hypothesis, shows that the relationship between earnings value-added and non-STEM value-added becomes significantly more positive when controlling for STEM value-added. On the other hand, STEM value-added becomes an even stronger predictor of earnings value-added when controlling for non-STEM value-added (bottom left panel of Figure 12). The partial slope coefficients imply that a 10 percentage point controlled increase in STEM value-added predicts roughly twice as much earnings value-added (\$3,540) as the same controlled increase in non-STEM value-added (\$1,739).

Figure 12: Value-Added on Earnings vs. Value-Added on STEM and Non-STEM BA Completion



Notes: Empirical Bayes shrunk value-added estimates are used as regressors to correct for mild attenuation bias from estimation error in value-added (see Appendix A.1). The top two panels plot earnings value-added against EB-shrunk STEM BA (left) and non-STEM BA (right) completion value-added. The middle panel plots value-added on STEM BA completion against value-added on non-STEM BA completion. The bottom left panel plots earnings value-added against the residual from a college-level regression of EB-shrunk STEM BA value-added on EB-shrunk non-STEM BA value-added. The bottom right panel plots earnings value-added against the residual from a college-level regression of EB-shrunk non-STEM BA value-added on EB-shrunk STEM BA value-added. Correlations, regressions, and circles are weighted by student enrollment.

### 7.3 Other Potential Mechanisms: Persistence, Transfer, and Industry

Figure 13 highlights the predictive power of other intermediate outcomes on earnings value-added. The top left panel shows that value-added on years of college completed, a more continuous measure of educational attainment, exhibits a strong correlation with earnings value-added: an extra year of attainment value-added is associated with an additional \$7,872 in earnings value-added.<sup>37</sup> Relatedly, the top right panel shows that a college's value-added on years of college completed is a nearly perfect predictor of its value-added on BA completion; this is far from obvious given that over a third of college-goers never complete a degree. The middle two panels illustrate transfer as a potentially detrimental mediator: colleges that induce more students to transfer to other institutions (horizontal axis) also tend to have lower value added on earnings (left) and degree completion (right). Finally, the bottom two panels illustrate the potential for college-industry linkages to help explain differences in earnings effects. As a test case, the bottom left plot shows a strong positive correlation between a college's value-added on working in the oil and gas industry and its earnings value-added. The bottom right plot, which shows the correlation between value-added on oil and gas employment and STEM degrees, suggests that majors may play a role in these linkages.

## 8 Match Effects: Student Heterogeneity in College Value-Added

Our final set of results probe the potential for (mis)match effects by allowing the impacts of attending different colleges to vary flexibly across our suite of observable student characteristics. We first visualize separate value-added estimates at each college for different student subpopulations to gauge the overall quantitative importance of match effects, then analyze student heterogeneity in the returns to selectivity to speak directly to empirical questions that arise in debates over the role of mismatch in higher education.

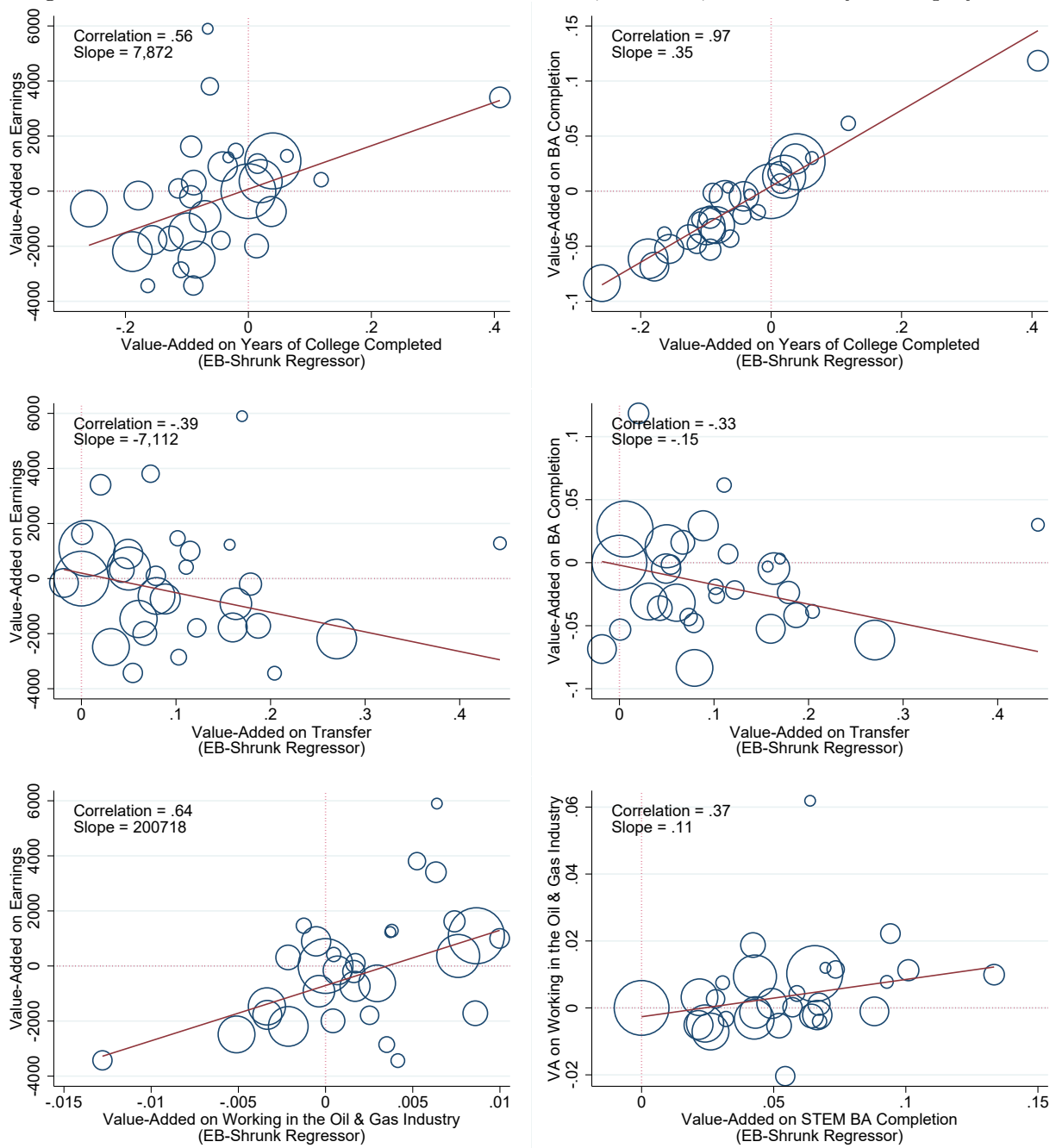
### 8.1 Value-Added Estimates Across Student Subpopulations

Figure 14 plots the distribution of value-added estimates conditional on different student subpopulations. Following the graphical format from Section 4, we plot the raw outcome means at each college for comparison, and our main value-added estimates (which pool across all student subgroups) appear in solid circles. Below each college's main estimate is the set of subgroup value-added estimates that come from estimating

---

<sup>37</sup>Under the strong exclusion restriction that different colleges only affect student earnings through their differential effects on persistence, this slope could be interpreted as an IV estimate of the return to a year of college, with the college indicators acting as instruments, the horizontal axis as the first stage, and the vertical axis as the reduced form. \$7,872 is a gain of 17.6% above the sample mean of \$44,834, which is on the high end of the distribution of such estimated returns (e.g. Card, 2001; Oreopoulos and Petronijevic, 2013), suggesting violations of exclusion by colleges adding value through other channels besides persistence.

Figure 13: Other Potential Mechanisms: Persistence, Transfer, and Industry of Employment



Notes: Empirical Bayes shrunk value-added estimates are used as regressors to correct for mild attenuation bias from estimation error in value-added (see Appendix A.1). Each panel plots our main value-added estimate on one outcome against the EB-shrunk value-added estimate on another outcome. Years of college completed, transfer, and working in the oil and gas industry are defined in Section 2. Correlations, regressions, and circles are weighted by student enrollment.

our main specification within each split of the sample: men vs. women, high income (non-FRPL) vs. low income (FRPL), high test score (above median) vs. low test score (below median), high attendance (above median) vs. low attendance (below median), and white vs. Hispanic vs. Black.

Two patterns emerge immediately from this exercise. First, a handful of estimates are extremely imprecise, with large confidence intervals and outlier point estimates. These come from college-subpopulation cells with small numbers of students, e.g., Black students at TAMU-International, which is 95% Hispanic, and white students at Prairie View A&M, which is 96% Black.<sup>38</sup> Second, and more substantively, we see that the value-added estimates across student subgroups tend to cluster around the main pooled estimate for each college, with deviations rarely statistically distinguishable from each other.<sup>39</sup>

To quantify the role of student-college match effects in explaining student outcomes, we also run a pooled specification in the full sample, reported in Tables B.7 and B.8, that interacts each college treatment (and each admissions portfolio) with each student covariate: gender, low-income (FRPL), race, standardized test score, and standardized attendance. An F-test shows that these treatment-covariate interactions are jointly statistically significant at the 1% level, suggesting match effects may be present, but the question remains as to whether they are economically meaningful. To gauge the magnitude of this heterogeneity in driving student outcomes, we can compare the  $R^2$  of this fully interacted specification to a specification without treatment-covariate interactions (but keeping admission portfolio-covariate interactions to maintain a common control set). This comparison attributes only a trivial contribution of match effects to explaining student outcomes, with the  $R^2$  increasing from .2423 to .2428 for BA completion (.2088 to .2090 in adjusted  $R^2$ ), and from .1611 to .1616 for earnings (.1206 to .1207 in adjusted  $R^2$ ).

These results suggest a limited scope for student heterogeneity in college value-added, at least across the dimensions of diversity that we can observe in our data. Methodologically, this may also help to allay concerns that pooling the estimation of college value-added across different types of students introduces biases from match effects. We provide further evidence of this in Figure A.5, which shows that our main estimates from Section 4 are quite similar to estimates from the specification with treatment-covariate interactions evaluated for a student with average covariate values.<sup>40</sup>

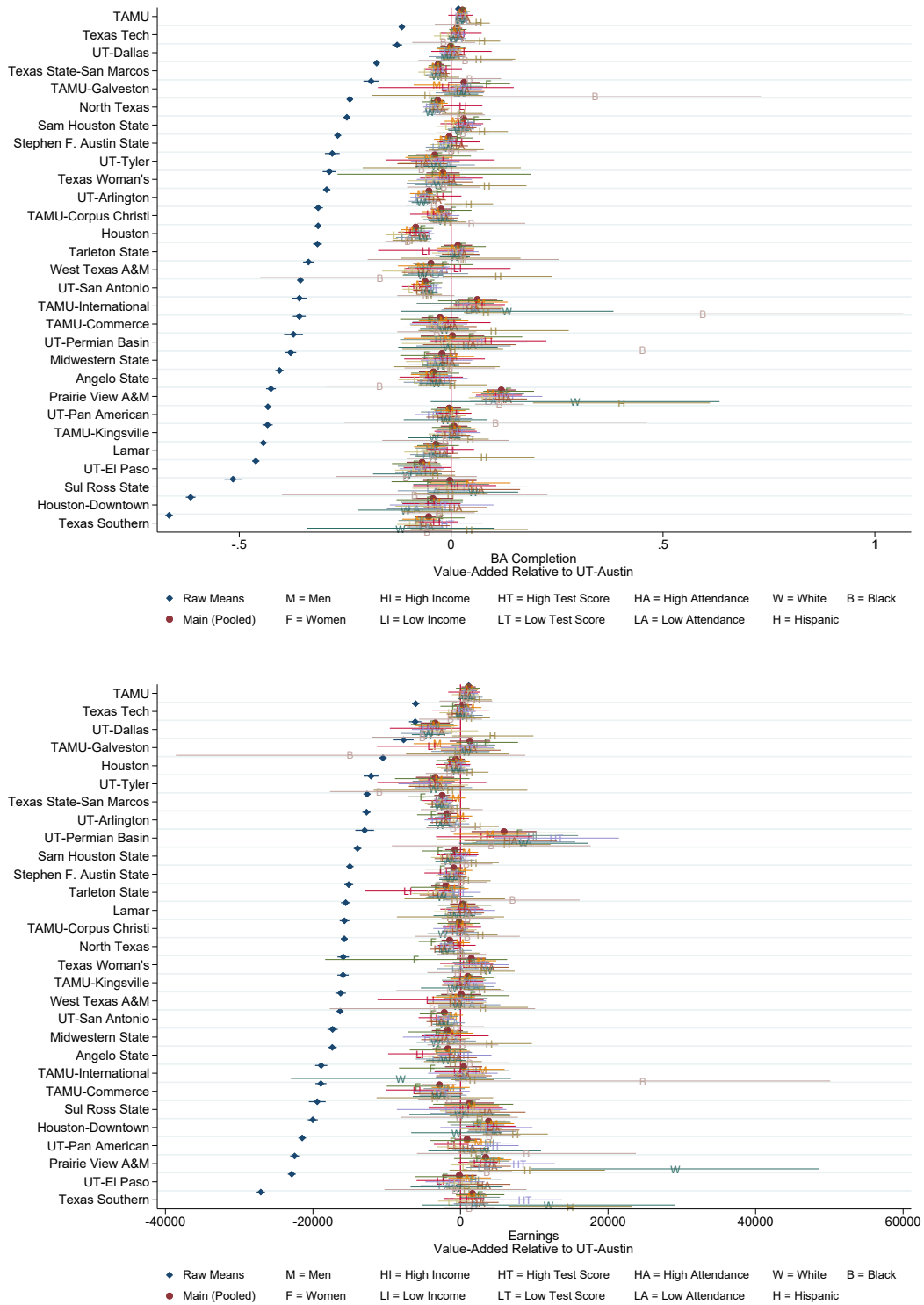
---

<sup>38</sup>All of these cells have at least 5 students in them; the smallest is 19 Black students at TAMU-International.

<sup>39</sup>One potentially interesting observation for future research is that Hispanics appear to have somewhat higher value-added estimates at some schools relative to UT-Austin, more so for BA completion than earnings, but it is not obvious what these schools have in common and thus whether this is a systematic pattern or idiosyncratic variation.

<sup>40</sup>We drop the five colleges that enroll almost exclusively one particular demographic group—Blacks at Prairie View A&M and Texas Southern, Hispanics at TAMU-International and UT-Pan American, and women at Texas Woman’s University—since these have extremely imprecise estimates for the average student. The two schools with non-trivial deviations between the two estimates—TAMU-Galveston and UT-Permian Basin—are two of our smallest schools, and thus their value-added estimates also become noisier when interacted with covariates.

Figure 14: Value-Added Estimates Across Student Subpopulations



Notes: Each set of point estimates and robust 95% confidence intervals come from regressions of individual student outcomes on college treatment indicators, omitting UT-Austin as the reference treatment (signified by the vertical line at zero). All specifications control for cohort fixed effects. The Raw Means specification controls for nothing else. The Main Specification adds our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. Each remaining specification estimates the Main Specification in a specified subpopulation of students: men vs. women, high income (non-FRPL) vs. low income (FRPL), high 10th grade test score (above median) vs. low 10th grade test score (below median), high high school attendance (above median) vs. low high school attendance (below median), and white vs. Hispanic vs. Black.

## 8.2 Testing for Mismatch: Heterogeneity in the Returns to Selectivity

Finally, our results in Figure 15 directly examine whether the returns to attending a more selective college vary systematically across different types of students. Each estimate is the coefficient from a student-level regression of a student’s outcome on the mean SAT score of her college peers. For comparison, the estimate at the top of each plot includes no controls, and the next is the overall pooled estimate from our main specification, which controls for admissions portfolio fixed effects and our core set of student covariates. We then run this specification separately for each student subpopulation, as in the previous subsection: men vs. women, high income vs. low income, high test score vs. low test score, high attendance vs. low attendance, and white vs. Hispanic vs. Black.

The results in Figure 15 show that for each outcome, the estimated impacts of attending a more selective college are rather similar across student subpopulations. We find little evidence of significant heterogeneity in the returns to selectivity by gender, family income as measured by FRPL eligibility, cognitive skills as measured by high school test scores, and non-cognitive skills as measured by high school attendance.<sup>41</sup> Black students are the main exception to this trend: they experience economically small but statistically significant negative impacts of selectivity, with a 100-point increase in peer mean SAT scores predicting a 1 percentage point reduction in BA completion, a 3 percentage point reduction in STEM completion (compared to a 2 percentage point reduction for the average student), and a \$1,000 reduction in earnings.

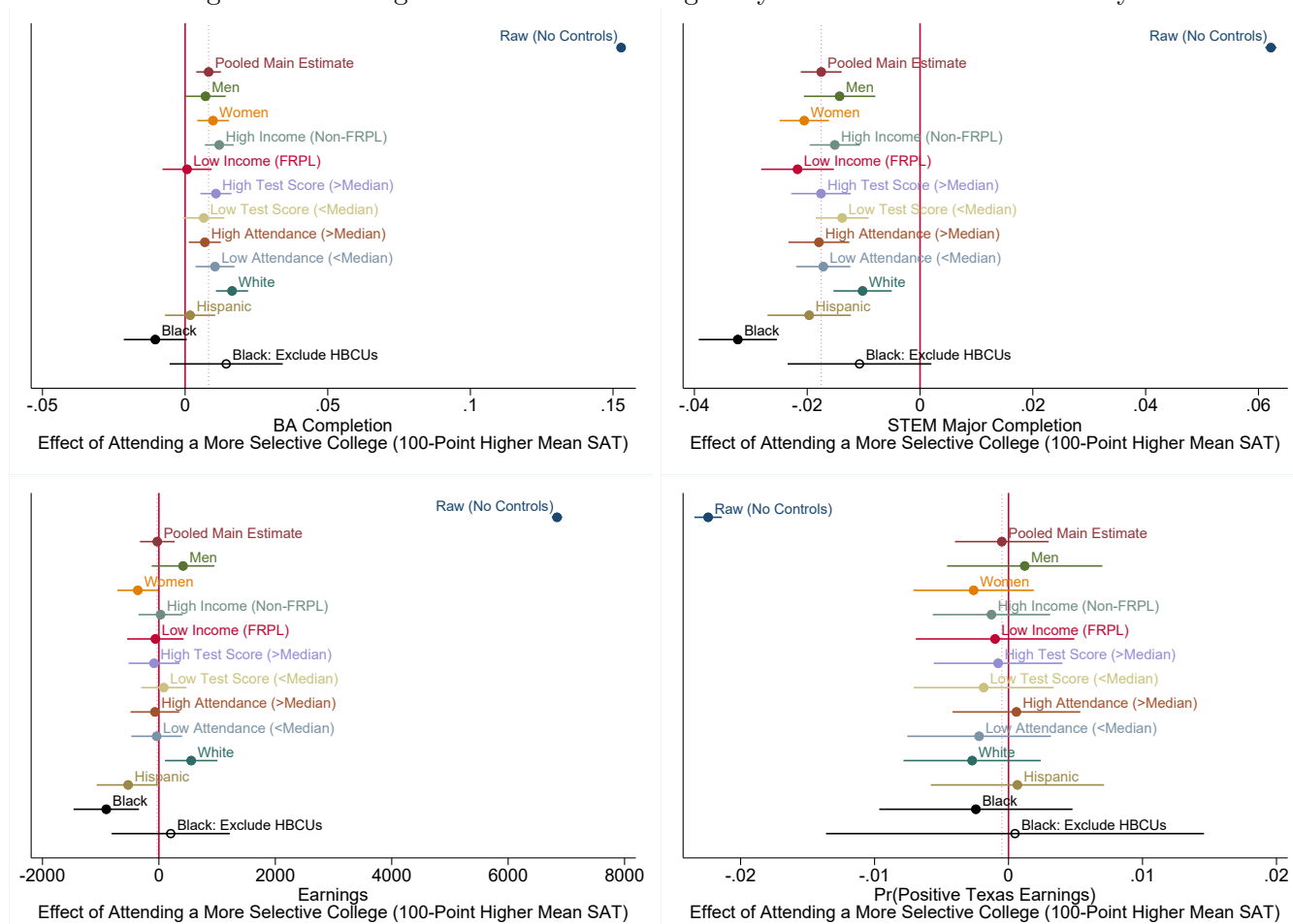
The estimates at the bottom of each plot demonstrate that this modest pattern of “mismatch” for Black students is driven by the two large historically Black universities (HBCUs) in Texas: Prairie View A&M and Texas Southern. The selectivity coefficients for Black students increase to become indistinguishable from the pooled estimate when we exclude these two schools from the sample. These two HBCUs have the lowest mean SAT scores among Texas public universities, but they generate above-average value-added for their students—especially Prairie View A&M—as shown in Figure 5. Their presence thus exerts downward pressure on the association between selectivity and value-added for Black students, as Black students comprise over 95% of their enrollment, and 28% of Black students in our sample attend one of them. Hence, the slight apparent “mismatch” effect for Black students is actually driven by their distinct college choice set relative to other students, rather than by significant heterogeneity in the effects of attending a given selective college. That is, across the non-HBCUs, i.e. the schools that non-Black students attend, Black students experience very similar impacts of selectivity as their peers from other backgrounds.

---

<sup>41</sup>We also find no gender difference in the effect of selectivity on appearing in the Texas earnings data, as shown in the bottom right panel of Figure 15. This contrasts with some of the results in Ge et al. (2018), though the difference may be driven by the fact that the women in their sample, the College and Beyond survey, attended college in the 1970s, compared to our cohorts enrolling 30 years later in the 2000s.



Figure 15: Testing for Mismatch: Heterogeneity in the Returns to Selectivity



Notes: Each point estimate and robust 95% confidence interval comes from a regression of individual student outcomes on the mean incoming SAT score of the student's college. The coefficients are scaled to correspond to a 100-point increase in mean SAT scores. All specifications control for cohort fixed effects. The Raw specification controls for nothing else. The Pooled Main Estimate adds college admission portfolio fixed effects and our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. Each remaining specification estimates the main specification in a specified subpopulation of students: men vs. women, high income (non-FRPL) vs. low income (FRPL), high 10th grade test score (above median) vs. low 10th grade test score (below median), high high school attendance (above median) vs. low high school attendance (below median), and white vs. Hispanic vs. Black. The final specification in each plot omits the two historically Black universities, Texas Southern and Prairie View A&M, from the set of college treatments in the subsample of Black students.

## 9 Conclusion

Average student outcomes vary enormously across colleges in the United States. We find that these outcome differences attenuate dramatically when comparing students who applied to and gained admission to the same set of schools, suggesting that value-added differences across the colleges in student choice sets play a smaller role than systematic student sorting in producing these observed outcome disparities. The distribution of college value-added is not degenerate, however, and while it is poorly proxied by selectivity, we find that non-peer college inputs like instructional spending do covary positively with value-added, especially conditional on selectivity, though residual value-added differences remain that are not neatly summarized by observable quality measures. Colleges that boost earnings also tend to boost intermediate outcomes like persistence and BA completion, shedding light on the potential mechanisms through which colleges shape the path from enrollment to earnings. We also examine whether college value-added varies systematically across student characteristics; we fail to find evidence of substantial match effects for most groups, with a small selectivity mismatch effect for Black students driven by a different choice set (the availability of HBCUs) rather than heterogeneity in value-added schedules across the same schools relative to peers from other backgrounds.

While these results offer new insights into the consequences of college choices, the limitations of our approach point to fruitful avenues for future work. First, our data span a very large state with a rich diversity of students and a wide array of colleges, but our sample is ultimately limited to public universities in Texas. Determining whether the patterns we find extend to other settings and to other postsecondary sectors, including non-profit and for-profit private colleges, will be valuable for understanding variation in college value-added across other segments of the higher education landscape. Second, our investigation of match effects between colleges and students is limited to observable dimensions of heterogeneity; questions remain as to what extent students sort and gain heterogeneously along unobservable dimensions. Relatedly, our identification strategy isolates variation among students who applied to and were admitted to the same colleges, leaving open the possibility that college impacts could differ for students outside of these comparison sets, e.g. disadvantaged students induced to attend highly selective colleges through interventions that substantially change their application and enrollment behavior. Finally, while our data linkages permit the measurement of college value-added on multiple policy-relevant outcomes, a wealth of other important outcomes remain to explore, including family formation (marriage, spousal characteristics, spousal income, childbearing), occupational choices, employer characteristics, and geographic mobility.

## References

- ABDULKADIROGLU, A., P. A. PATHAK, J. SCHELLENBERG, AND C. R. WALTERS (2020): “Do Parents Value School Effectiveness?” *American Economic Review*, 110, 1502–1539.
- ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): “High Wage Workers and High Wage Firms,” *Econometrica*, 67, 251–333.
- AISCH, G., R. GEBELOFF, AND K. QUEALY (2014): “Where We Came From and Where We Went, State by State,” *New York Times*, 19 August.
- ALTONJI, J. G. AND C. R. PIERRET (2001): “Employer Learning and Statistical Discrimination,” *The Quarterly Journal of Economics*, 116, 313–350.
- ANDREWS, R. J., J. LI, AND M. F. LOVENHEIM (2014): “Heterogeneous Paths through College: Detailed Patterns and Relationships with Graduation and Earnings,” *Economics of Education Review*, 42, 93–108.
- (2016): “Quantile Treatment Effects of College Quality on Earnings,” *Journal of Human Resources*, 51, 200–238.
- ANELLI, M. (2019): “Returns to Elite University Education: A Quasi-experimental Analysis,” *Journal of European Economic Association*, forthcoming.
- ANGRIST, J. D. AND S. H. CHEN (2011): “Schooling and the Vietnam-Era GI Bill: Evidence from the Draft Lottery,” *American Economic Journal: Applied Economics*, 3, 96–118.
- ANGRIST, J. D., P. D. HULL, P. A. PATHAK, AND C. R. WALTERS (2017): “Leveraging Lotteries for School Value-Added: Testing and Estimation,” *The Quarterly Journal of Economics*, 132, 871–919.
- ANGRIST, J. D., P. A. PATHAK, AND R. A. ZARATE (2019): “Choice and Consequence: Assessing Mismatch at Chicago Exam Schools,” NBER Working Paper No. 26137.
- ARCIDIACONO, P. (2005): “Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?” *Econometrica*, 73, 1477–1524.
- ARCIDIACONO, P., E. M. AUCEJO, AND V. J. HOTZ (2016): “University Differences in the Graduation of Minorities in STEM Fields: Evidence from California,” *American Economic Review*, 106, 525–562.
- ARCIDIACONO, P. AND M. LOVENHEIM (2016): “Affirmative Action and the Quality-Fit Trade-off,” *Journal of Economic Literature*.

- ARMONA, L., R. CHAKRABARTI, AND M. F. LOVENHEIM (2018): “How Does For-Profit College Attendance Affect Student Loans, Defaults and Labor Market Outcomes?” NBER Working Paper No. 25042.
- BLACK, D. A. AND J. A. SMITH (2006): “Estimating the Returns to College Quality with Multiple Proxies for Quality,” *Journal of Labor Economics*, 24, 701–728.
- BLACK, S. E., J. T. DENNING, AND J. ROTHSTEIN (2020): “Winners and Losers? The Effect of Gaining and Losing Access to Selective Colleges on Education and Labor Market Outcomes,” NBER Working Paper No. 26821.
- BODOH-CREED, A. L. AND B. R. HICKMAN (2019): “Pre-College Human Capital Investment and Affirmative Action: A Structural Policy Analysis of US College Admissions,” Working paper, Washington University in St. Louis.
- BOWEN, W. G. AND D. C. BOK (2000): *The Shape of the River: Long-term Consequences of Considering Race in College and University Admissions*, Book collections on Project MUSE, Princeton University Press.
- BREWER, D. J. AND R. G. EHRENBERG (1996): “Does It Pay to Attend an Elite Private College? Evidence from the Senior High School Class of 1980,” *Research in Labor Economics*, 15, 239–271.
- BREWER, D. J., E. R. EIDE, AND R. G. EHRENBERG (1999): “Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings,” *The Journal of Human Resources*.
- BROECKE, S. (2012): “University selectivity and earnings: Evidence from UK data on applications and admissions to university,” *Economics of Education Review*, 31, 96–107.
- CANAAN, S. AND P. MOUGANIE (2018): “Returns to Education Quality for Low-Skilled Students: Evidence from a Discontinuity,” *Journal of Labor Economics*, 36, 395–436.
- CARD, D. (2001): “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems,” *Econometrica*.
- CARNEIRO, P., J. J. HECKMAN, AND E. J. VYTLACIL (2011): “Estimating Marginal Returns to Education,” *American Economic Review*, 101, 2754–2781.

- CELLINI, S. R. AND N. TURNER (2019): “Gainfully Employed?: Assessing the Employment and Earnings of For-Profit College Students Using Administrative Data,” *Journal of Human Resources*, 54, 342–370.
- CHETTY, R., J. N. FRIEDMAN, AND J. E. ROCKOFF (2014a): “Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates,” *American Economic Review*, 104, 2593–2632.
- (2014b): “Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood,” *American Economic Review*, 104, 2633–2679.
- CHETTY, R., J. N. FRIEDMAN, E. SAEZ, N. TURNER, AND D. YAGAN (2020): “Income Segregation and Intergenerational Mobility Across Colleges in the United States,” *The Quarterly Journal of Economics*, 135, 1567–1633.
- CHETTY, R. AND N. HENDREN (2018): “The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates,” *The Quarterly Journal of Economics*, 133, 1163–1228.
- CHETTY, R., N. HENDREN, P. KLINE, AND E. SAEZ (2014c): “Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States,” *The Quarterly Journal of Economics*, 129, 1553–1623.
- COHODES, S. R. AND J. S. GOODMAN (2014): “Merit Aid, College Quality, and College Completion: Massachusetts’ Adams Scholarship as an In-Kind Subsidy,” *American Economic Journal: Applied Economics*, 6, 251–285.
- CUNHA, J. M. AND T. MILLER (2014): “Measuring Value-Added in Higher Education: Possibilities and Limitations in the Use of Administrative Data,” *Economics of Education Review*, 42, 64–77.
- DALE, S. B. AND A. B. KRUEGER (2002): “Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables,” *The Quarterly Journal of Economics*, 117, 1491–1527.
- (2014): “Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data,” *Journal of Human Resources*, 49, 323–358.
- DILLON, E. W. AND J. SMITH (2018): “The Consequences of Academic Match between Students and Colleges,” NBER Working Paper No. 25069.

- DYNARSKI, S., C. LIBASSI, K. MICHELMORE, AND S. OWEN (2020): “Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students,” NBER Working Paper No. 25349.
- ELLER, C. C. (2019): “Superficially Coupled Systems: The Organizational Production of Inequality in Higher Education,” Annenberg EdWorkingPaper No. 19–70.
- FARBER, H. S. AND R. GIBBONS (1996): “Learning and Wage Dynamics,” *The Quarterly Journal of Economics*, 111, 1007–1047.
- FINKELSTEIN, A., M. GENTZKOW, AND H. WILLIAMS (2016): “Sources of Geographic Variation in Health Care: Evidence From Patient Migration,” *The Quarterly Journal of Economics*, 131, 1681–1726.
- FRYER, R. G. AND M. GREENSTONE (2010): “The Changing Consequences of Attending Historically Black Colleges and Universities,” *American Economic Journal: Applied Economics*, 2, 116–148.
- GE, S., E. ISAAC, AND A. MILLER (2018): “Elite Schools and Opting-In: Effects of College Selectivity on Career and Family Outcomes,” NBER Working Paper No. 25315.
- GOODMAN, J., M. HURWITZ, AND J. SMITH (2017): “Access to Four-Year Public Colleges and Degree Completion,” *Journal of Labor Economics*, 35, 829–867.
- HASTINGS, J. S., C. A. NEILSON, AND S. D. ZIMMERMAN (2014): “Are Some Degrees Worth More than Others? Evidence from College Admission Cutoffs in Chile,” NBER Working Paper No. 19241.
- HECKMAN, J. J., J. E. HUMPHRIES, AND G. VERAMENDI (2018): “Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking,” *Journal of Political Economy*, 126, S197–S246.
- HOEKSTRA, M. (2009): “The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach,” *Review of Economics and Statistics*, 91, 717–724.
- HOXBY, C. AND S. TURNER (2013): “Expanding College Opportunities for High-Achieving, Low Income Students,” Tech. Rep. 12-014.
- HOXBY, C. M. (2019): “The Productivity of US Postsecondary Institutions,” in *Productivity in Higher Education*, ed. by C. M. Hoxby and K. Stange, University of Chicago Press, 31–66.
- JACOB, B. AND L. LEFGREN (2008): “Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education,” *Journal of Labor Economics*, 26, 101–136.

- JIA, R. AND H. LI (2019): “Designated Elitism,” Working paper.
- KIRKEBOEN, L. J., E. LEUVEN, AND M. MOGSTAD (2016): “Field of Study, Earnings, and Self-Selection,” *The Quarterly Journal of Economics*, 131, 1057–1111.
- KRUEGER, A. B. AND L. H. SUMMERS (1988): “Efficiency Wages and the Inter-Industry Wage Structure,” *Econometrica*, 56, 259.
- LEMIEUX, T. AND D. CARD (2001): “Education, Earnings, and the Canadian G.I. Bill,” *Canadian Journal of Economics/Revue canadienne d’économique*, 34, 313–344.
- LONG, M. C. (2008): “College Quality and Early Adult Outcomes,” *Economics of Education Review*, 27, 588–602.
- MORRIS, C. N. (1983): “Parametric Empirical Bayes Inference: Theory and Applications,” *Journal of the American Statistical Association*, 78, 47–55.
- MOUNTJOY, J. (2019): “Community Colleges and Upward Mobility,” Working Paper.
- NATIONAL CENTER FOR EDUCATION STATISTICS (2018): “Digest of Education Statistics,” Tech. rep.
- NATIONAL CONFERENCE OF STATE LEGISLATURES (2018): “Outcomes-Based Funding as an Evolving State Appropriation Model,” Tech. rep.
- OREOPOULOS, P. AND U. PETRONIJEVIC (2013): “Making College Worth It: A Review of the Returns to Higher Education,” *The Future of Children*, 23, 41–65.
- REYNOLDS, C. L. (2012): “Where to Attend? Estimating the Effects of Beginning College at a Two-Year Institution,” *Economics of Education Review*, 31, 345–362.
- ROBBINS, H. (1956): “An Empirical Bayes Approach to Statistics,” in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 1*, ed. by J. Neyman, Berkeley: University of California Press, 157–163.
- SAAVEDRA, J. (2009): “The Learning and Early Labor Market Effects of College Quality: A Regression Discontinuity Analysis,” Working paper.
- SANDER, R. H. AND S. TAYLOR (2012): *Mismatch: How Affirmative Action Hurts Students It’s Intended to Help, and Why Universities Won’t Admit It*, Basic Books.

- SEKHRI, S. (2020): “Prestige Matters: Wage Premium and Value Addition in Elite Colleges,” *American Economic Journal: Applied Economics*, 12, 207–225.
- SMITH, J., J. GOODMAN, AND M. HURWITZ (2020): “The Economic Impact of Access to Public Four-Year Colleges,” NBER Working Paper No. 27177.
- STEVENS, D. W. (2007): “Employment That Is Not Covered by State Unemployment Insurance Laws,” U.S. Census Bureau Technical Paper No. TP–2007–04.
- ZIMMERMAN, S. D. (2014): “The Returns to College Admission for Academically Marginal Students,” *Journal of Labor Economics*, 32, 711–754.



# A Appendix

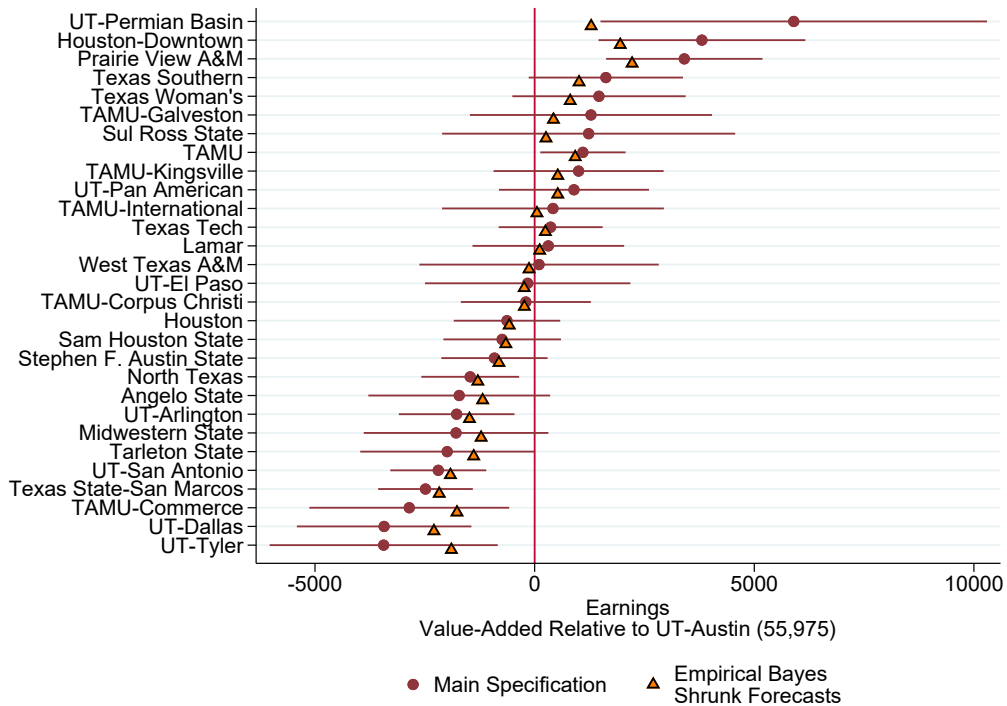
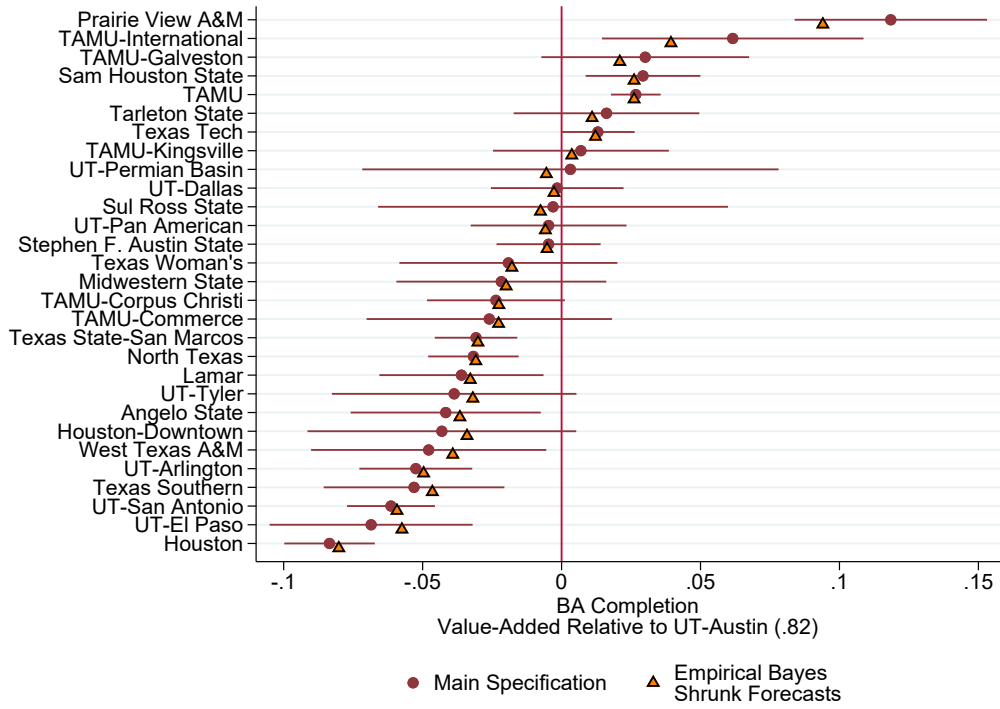
## A.1 Empirical Bayes Shrunk Forecasts

As a means of accounting for estimation error in value-added, we construct Empirical Bayes forecasts (e.g., Robbins, 1956) that shrink each college’s value-added estimate towards the mean in proportion to its imprecision. These EB estimates are interesting in themselves as efficient predictions of true college value-added given our finite data, but they deliver additional utility as error-correcting regressors (e.g., Jacob and Lefgren, 2008) in the analysis in Section 7.

Following the standard approach in the value-added literature, we can derive the EB shrinkage factor either by assuming a normally distributed prior over true value-added  $v_j$  and estimation error  $e_j$  (e.g., Morris, 1983; Abdulkadiroglu et al., 2020), or by restricting our forecast function of  $v_j$  given  $\hat{v}_j$  to linear projection (e.g., Chetty et al., 2014a). Both premises deliver a shrinkage factor of the form  $\frac{\sigma_v^2}{\sigma_v^2 + \sigma_{e_j}^2}$ , where  $\sigma_{e_j}^2$  is the college  $j$ -specific estimation error variance. We estimate  $\sigma_v^2$  using the empirical counterparts of Equation (7), and  $\sigma_{e_j}^2$  is estimated as the square of the standard error of our value-added estimate  $\hat{v}_j$  for college  $j$ .

Figure A.1 presents the Empirical Bayes forecasts and compares them to the unshrunk value-added estimates. Since the value-added estimates for most of our colleges are measured relatively precisely and/or are already close to zero, the EB shrinkage procedure does not have much impact on them; as expected, the greatest shrinkage is exerted on the handful of small schools with noisier and more far-flung estimates at the top and bottom of the distribution.

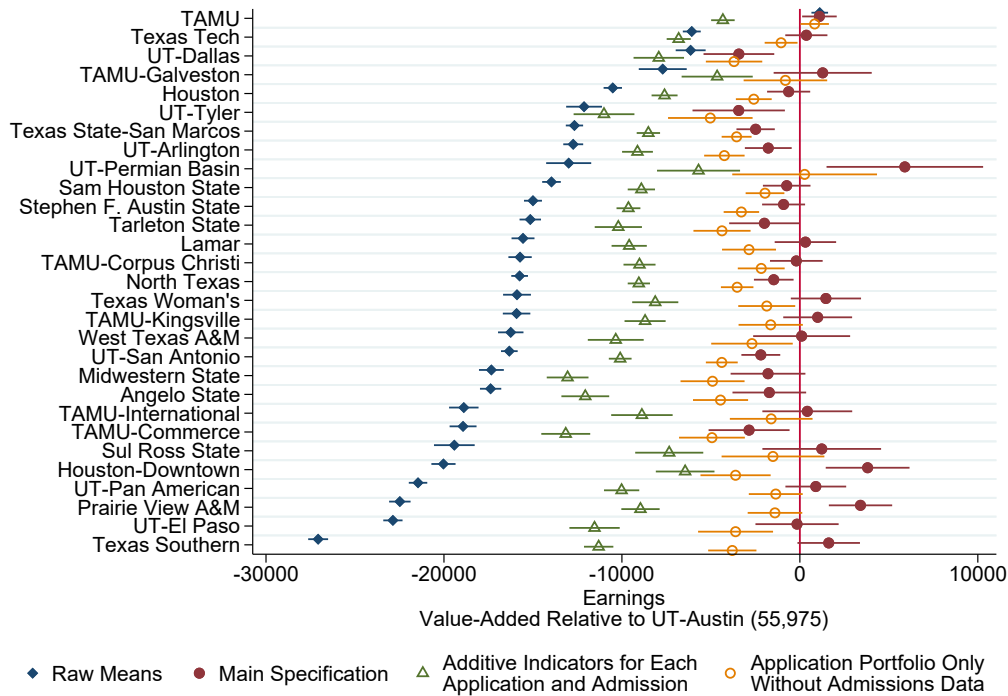
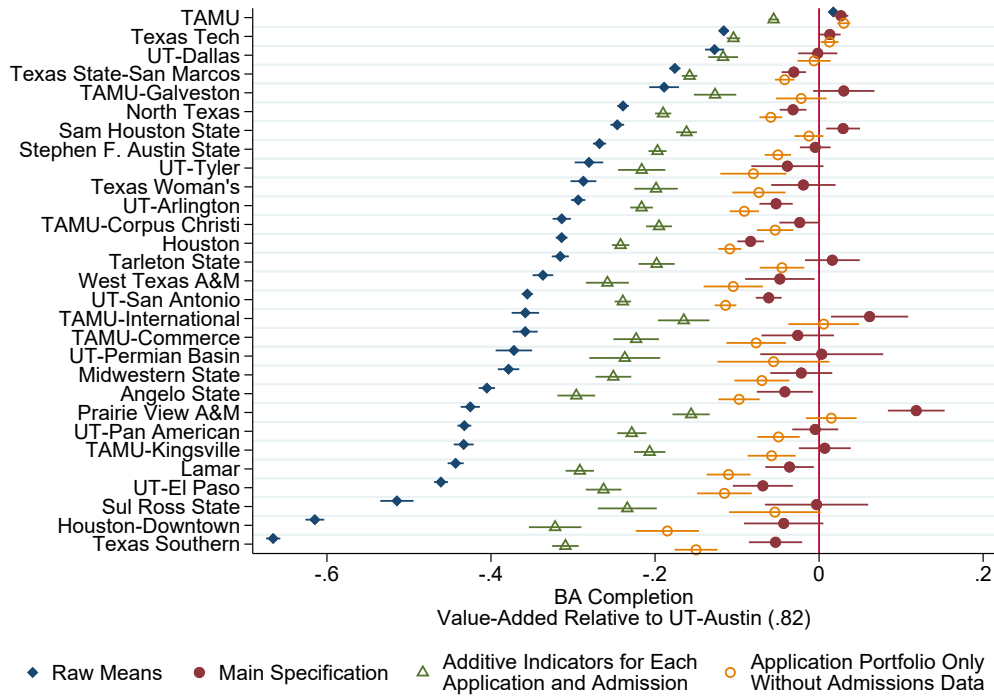
Figure A.1: Empirical Bayes Shrunk Forecasts of College Value-Added



Notes: The Main Specification point estimates and robust 95% confidence intervals come from regressions of individual student outcomes on college treatment indicators (with UT-Austin as the reference treatment at zero), controlling for admission portfolio fixed effects and our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. The Empirical Bayes Shrunk Forecasts shrink the Main Specification estimates towards their mean by the shrinkage factor  $\frac{\sigma_v^2}{\sigma_v^2 + \sigma_{\epsilon_j}^2}$ , as described in Section A.1.

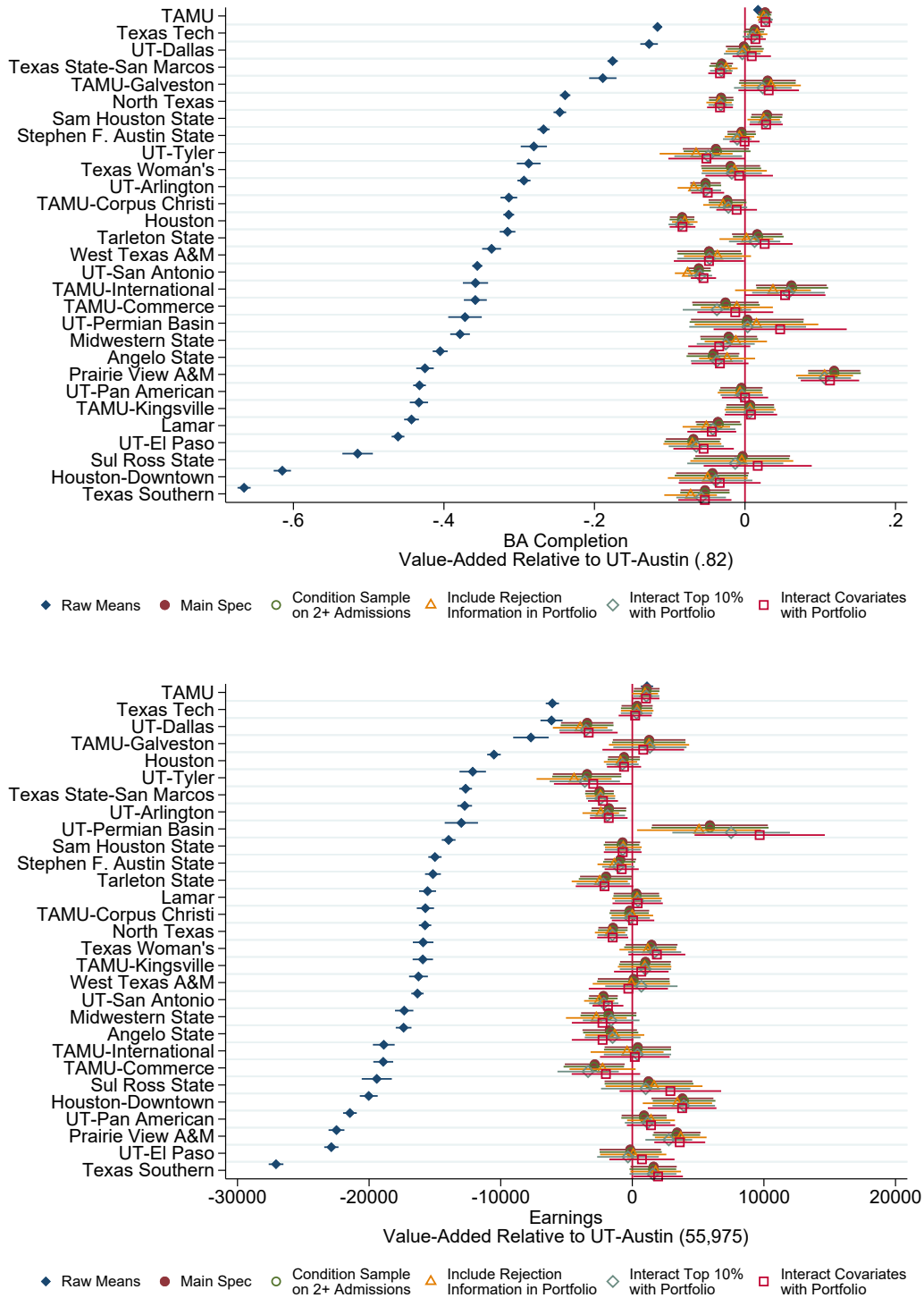
## A.2 Additional Figures and Tables

Figure A.2: Insufficiency of Simpler Specifications of Admissions Portfolios



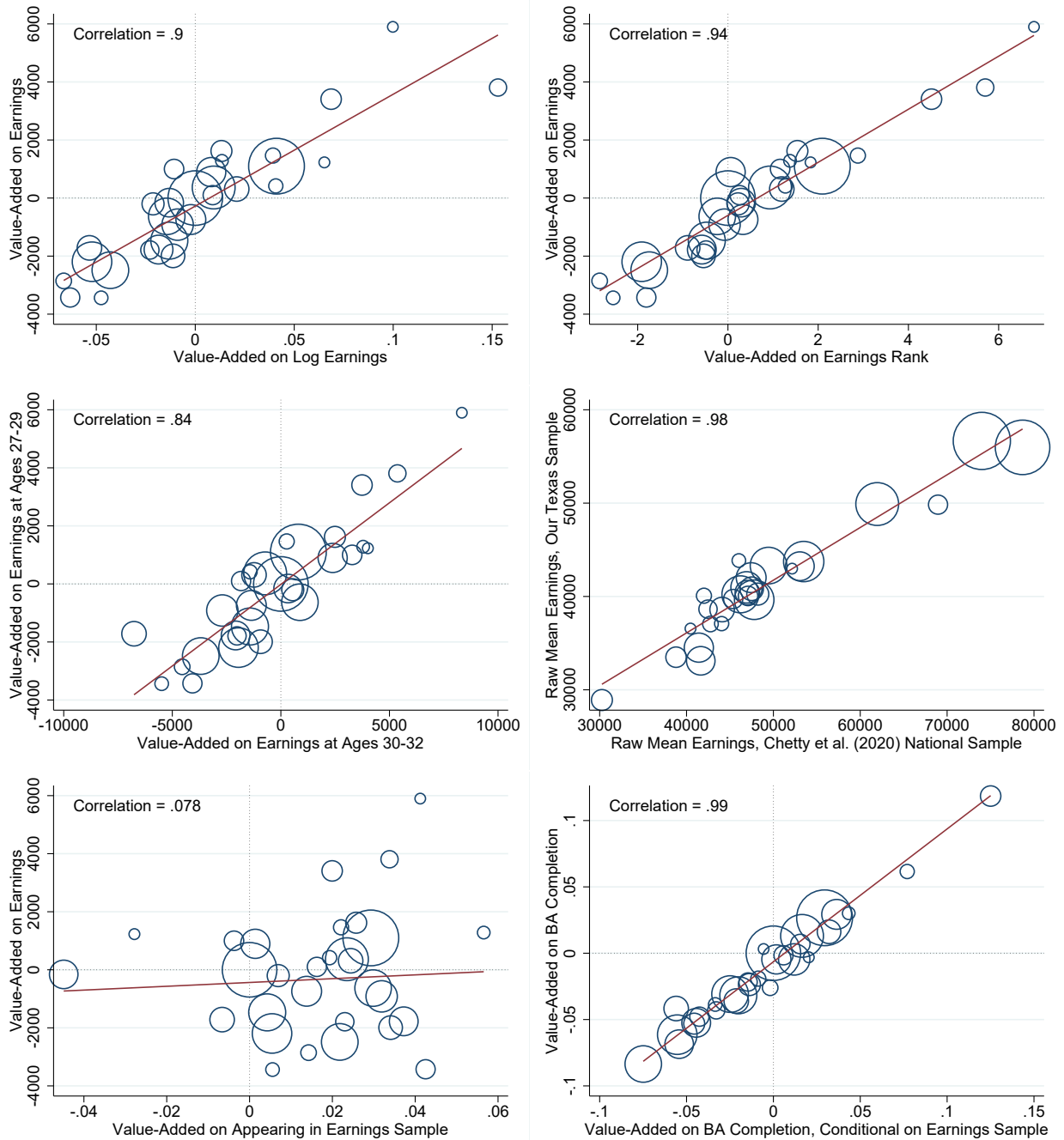
Notes: Each set of point estimates and robust 95% confidence intervals come from regressions of individual student outcomes on college treatment indicators, omitting UT-Austin as the reference treatment (signified by the vertical line at zero). The UT-Austin outcome mean appears in parentheses below each plot. All specifications control for cohort fixed effects. The *Raw Means* specification controls for nothing else. The *Main Specification* controls for college admission portfolio fixed effects and our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. The *Additive Indicators* specification replaces the full set of college admission portfolio fixed effects from the Main Specification with 29 dummies indicating an application to each college, and 29 dummies indicating an admission to each college, with UT-Austin as the omitted category. The *Application Portfolio Only* specification replaces the college admission portfolio fixed effects from the Main Specification with fixed effects solely for each combination of applications, ignoring information on admissions. See Appendix Tables B.4 and B.5 for the corresponding numerical estimates.

Figure A.3: Robustness to Richer Specifications of Admissions Portfolios



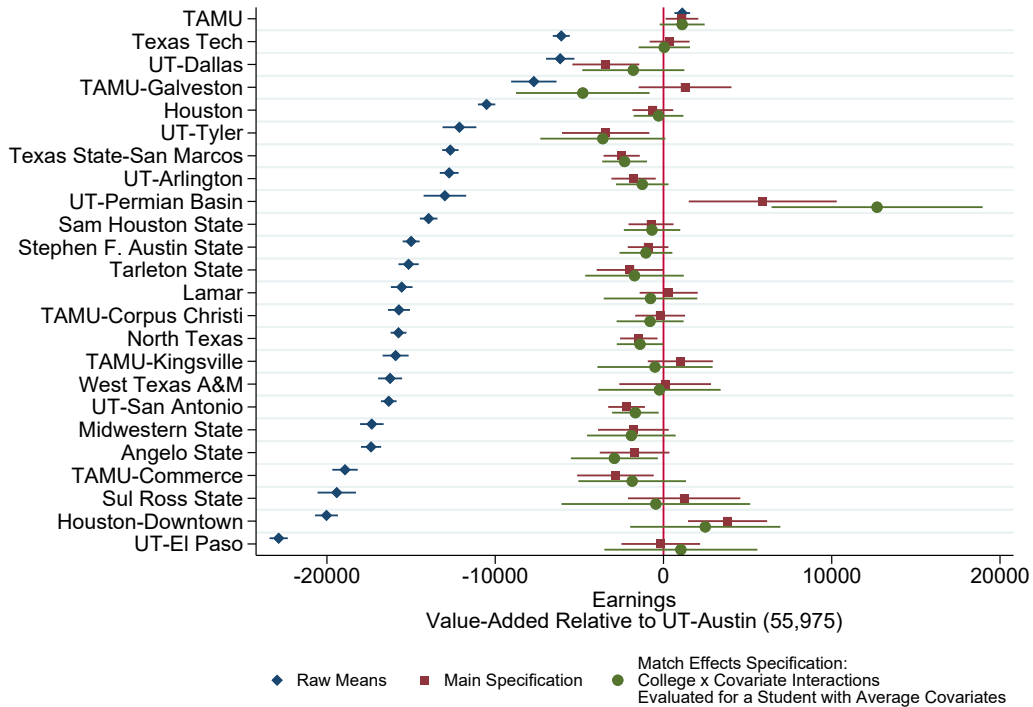
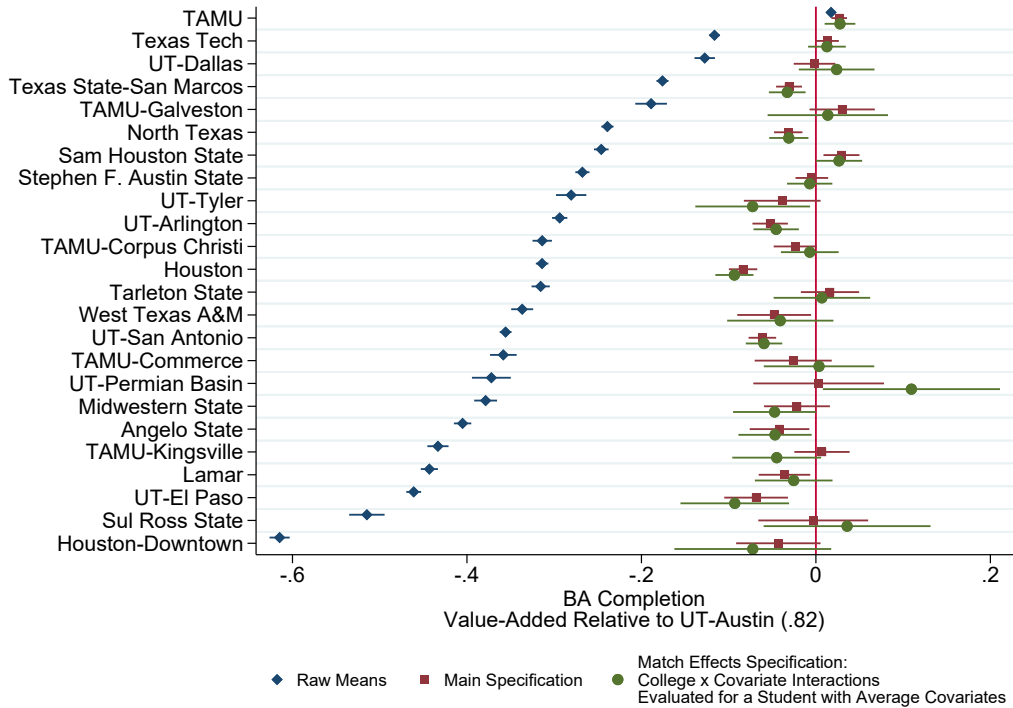
Notes: Each set of point estimates and robust 95% confidence intervals come from regressions of individual student outcomes on college treatment indicators, omitting UT-Austin as the reference treatment (signified by the vertical line at zero). The UT-Austin outcome mean appears in parentheses below each plot. All specifications control for cohort fixed effects. The *Raw Means* specification controls for nothing else. The *Main Specification* controls for college admission portfolio fixed effects and our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. The *Condition Sample* specification runs the Main Specification on the subsample of students who are admitted to at least two colleges. The *Include Rejection Information* specification replaces the admission portfolio fixed effects with fixed effects for every distinct portfolio of applications and admissions, thereby including additional information about applications that do not result in admissions, i.e. rejections. The *Interact Top 10%* specification interacts the admission portfolio indicators with the indicator for being in top decile of high school GPA. The *Interact Covariates* specification interacts the admission portfolio indicators with our core covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. See Appendix Tables B.4 and B.5 for the corresponding numerical estimates.

Figure A.4: Robustness to Alternative Earnings Definitions and Missing Earnings



Notes: Each circle plots college value-added estimates from our main specification, which regresses individual student outcomes on college treatment indicators (with UT-Austin as the reference treatment at zero), controlling for admission portfolio fixed effects and our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. The one exception is the fourth panel, which simply plots raw mean earnings at each college in our Texas sample against raw mean earnings from Chetty et al. (2020)'s national sample. All correlations and circles are weighted by student enrollment.

Figure A.5: Allowing for Match Effects Delivers Similar Estimates for the Average Student



Notes: Each set of point estimates and robust 95% confidence intervals come from regressions of individual student outcomes on college treatment indicators, omitting UT-Austin as the reference treatment (signified by the vertical line at zero). The UT-Austin outcome mean appears in parentheses below each plot. All specifications control for cohort fixed effects. The *Raw Means* specification controls for nothing else. The *Main Specification* adds college admission portfolio fixed effects and our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. The *Match Effects Specification* adds interactions between the college treatment indicators and the core covariate controls, and the value-added estimates are evaluated for a student with average values of the interacted covariates. See Tables B.7 and B.8 for the corresponding numerical estimates.

## B Online Appendix



Table B.1: Value-Added Estimates: BA Completion

	Raw Means	Typical Controls	Baseline	Main Spec	Extra Controls 1	Extra Controls 2
Angelo State	-0.405 (0.00508)	-0.232 (0.00494)	-0.0242 (0.0181)	-0.0417 (0.0175)	-0.0365 (0.0174)	-0.0274 (0.0170)
TAMU-Commerce	-0.358 (0.00779)	-0.186 (0.00745)	-0.0183 (0.0235)	-0.0260 (0.0225)	-0.0216 (0.0225)	-0.0266 (0.0222)
Lamar	-0.443 (0.00502)	-0.250 (0.00494)	-0.0329 (0.0156)	-0.0360 (0.0151)	-0.0333 (0.0150)	-0.0298 (0.0148)
Midwestern State	-0.379 (0.00668)	-0.215 (0.00640)	-0.00780 (0.0199)	-0.0217 (0.0193)	-0.0204 (0.0192)	-0.0146 (0.0191)
North Texas	-0.239 (0.00359)	-0.120 (0.00358)	-0.0208 (0.00855)	-0.0317 (0.00831)	-0.0288 (0.00830)	-0.0260 (0.00827)
UT-Pan American	-0.432 (0.00431)	-0.144 (0.00461)	-0.00492 (0.0148)	-0.00467 (0.0143)	-0.00133 (0.0142)	0.00153 (0.0141)
Sam Houston State	-0.246 (0.00428)	-0.0877 (0.00429)	0.0344 (0.0108)	0.0293 (0.0105)	0.0325 (0.0105)	0.0403 (0.0104)
Texas State-San Marcos	-0.176 (0.00353)	-0.0463 (0.00358)	-0.0384 (0.00778)	-0.0308 (0.00757)	-0.0254 (0.00757)	-0.0163 (0.00749)
Stephen F. Austin State	-0.268 (0.00413)	-0.116 (0.00412)	-0.00106 (0.00986)	-0.00467 (0.00955)	-0.00131 (0.00953)	0.00157 (0.00946)
Sul Ross State	-0.515 (0.0103)	-0.204 (0.0100)	-0.0167 (0.0336)	-0.00306 (0.0321)	0.00185 (0.0320)	0.000783 (0.0315)
Prairie View A&M	-0.425 (0.00594)	-0.0985 (0.00630)	0.122 (0.0183)	0.118 (0.0177)	0.120 (0.0176)	0.120 (0.0175)
Tarleton State	-0.315 (0.00532)	-0.162 (0.00526)	0.0370 (0.0176)	0.0162 (0.0170)	0.0199 (0.0170)	0.0232 (0.0169)
TAMU	0.0174 (0.00230)	0.0355 (0.00226)	0.0264 (0.00466)	0.0267 (0.00455)	0.0316 (0.00456)	0.0383 (0.00455)
TAMU-Kingsville	-0.433 (0.00619)	-0.164 (0.00605)	0.00755 (0.0169)	0.00696 (0.0161)	0.0117 (0.0161)	0.0147 (0.0159)
Texas Southern	-0.666 (0.00446)	-0.278 (0.00532)	-0.0568 (0.0169)	-0.0531 (0.0166)	-0.0492 (0.0165)	-0.0520 (0.0165)
Texas Tech	-0.116 (0.00304)	-0.0147 (0.00307)	0.0145 (0.00693)	0.0131 (0.00675)	0.0196 (0.00676)	0.0212 (0.00675)
Texas Woman's	-0.287 (0.00806)	-0.146 (0.00775)	-0.00180 (0.0206)	-0.0191 (0.0200)	-0.0163 (0.0199)	-0.0144 (0.0198)
Houston	-0.314 (0.00366)	-0.170 (0.00367)	-0.0829 (0.00852)	-0.0835 (0.00831)	-0.0811 (0.00830)	-0.0856 (0.00829)
UT-Arlington	-0.294 (0.00446)	-0.165 (0.00436)	-0.0498 (0.0107)	-0.0525 (0.0104)	-0.0524 (0.0103)	-0.0488 (0.0103)
UT-El Paso	-0.461 (0.00434)	-0.188 (0.00456)	-0.0791 (0.0191)	-0.0685 (0.0186)	-0.0673 (0.0186)	-0.0723 (0.0184)
West Texas A&M	-0.337 (0.00643)	-0.187 (0.00615)	-0.0342 (0.0224)	-0.0478 (0.0216)	-0.0425 (0.0215)	-0.0390 (0.0215)
TAMU-International	-0.358 (0.00855)	-0.112 (0.00829)	0.0759 (0.0250)	0.0616 (0.0240)	0.0561 (0.0239)	0.0599 (0.0239)
UT-Dallas	-0.127 (0.00596)	-0.0808 (0.00572)	0.00972 (0.0126)	-0.00157 (0.0122)	-0.00177 (0.0122)	-0.00571 (0.0121)
UT-Permian Basin	-0.372 (0.0114)	-0.190 (0.0109)	0.00152 (0.0394)	0.00317 (0.0382)	0.00491 (0.0381)	-0.00592 (0.0376)
UT-San Antonio	-0.356 (0.00344)	-0.158 (0.00355)	-0.0623 (0.00830)	-0.0614 (0.00806)	-0.0599 (0.00805)	-0.0570 (0.00797)
TAMU-Galveston	-0.189 (0.00927)	-0.0570 (0.00898)	0.0261 (0.0194)	0.0301 (0.0191)	0.0354 (0.0191)	0.0347 (0.0188)
TAMU-Corpus Christi	-0.314 (0.00567)	-0.135 (0.00554)	-0.0144 (0.0130)	-0.0236 (0.0127)	-0.0183 (0.0126)	-0.0113 (0.0124)
UT-Tyler	-0.281 (0.00889)	-0.183 (0.00851)	-0.0300 (0.0231)	-0.0386 (0.0224)	-0.0377 (0.0223)	-0.0366 (0.0219)
Houston-Downtown	-0.615 (0.00587)	-0.283 (0.00608)	-0.0507 (0.0252)	-0.0431 (0.0247)	-0.0431 (0.0246)	-0.0458 (0.0246)
R-Squared	0.1281	0.2130	0.1517	0.2166	0.2219	0.2443
N	422956	422956	418267	418267	418267	418166

Notes: These estimates correspond to Figures 2 and 4. Each column presents point estimates and robust standard errors from a regression of individual student outcomes on college treatment indicators, omitting UT-Austin as the reference treatment. The UT-Austin outcome mean is 0.82. All specifications control for cohort fixed effects. The Raw Means specification controls for nothing else. The Typical Controls specification adds controls for demographics (gender, race, FRPL), high school academic preparation (10th grade test scores, advanced coursework, and top high school GPA decile indicator), and behavioral measures of non-cognitive skills (high school attendance, disciplinary infractions, and an indicator for ever being at risk of dropping out). The Baseline Specification controls solely for college admission portfolio fixed effects (and cohort fixed effects). The Main Specification adds our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. Extra Control Set 1 adds controls for advanced high school coursework, disciplinary infractions, and an indicator for ever being at risk of dropping out. Extra Control Set 2 adds fixed effects for every high school and an indicator for being in the top decile of high school GPA.

Table B.2: Value-Added Estimates: Earnings

	Raw Means	Typical Controls	Baseline	Main Spec	Extra Controls 1	Extra Controls 2
Angelo State	-17377.9 (305.2)	-11312.6 (308.6)	-697.8 (1077.8)	-1716.1 (1056.2)	-1700.2 (1056.1)	-1197.0 (1051.3)
TAMU-Commerce	-18925.8 (385.9)	-12446.9 (385.1)	-2444.6 (1187.6)	-2854.1 (1160.7)	-2802.5 (1159.4)	-2905.0 (1155.1)
Lamar	-15556.9 (328.7)	-7641.1 (326.6)	1089.5 (906.6)	311.4 (880.8)	359.4 (879.4)	357.9 (881.7)
Midwestern State	-17334.2 (356.2)	-11589.6 (359.1)	-957.0 (1091.4)	-1789.6 (1072.3)	-1912.9 (1071.4)	-1575.1 (1076.6)
North Texas	-15747.0 (239.5)	-10861.6 (246.1)	-1227.4 (572.3)	-1467.1 (568.0)	-1484.0 (568.2)	-1389.4 (566.9)
UT-Pan American	-21456.6 (265.1)	-9832.4 (285.8)	470.8 (885.1)	895.5 (871.1)	888.1 (870.9)	1757.0 (869.9)
Sam Houston State	-13957.0 (267.1)	-7608.7 (273.8)	-851.0 (691.7)	-739.6 (682.4)	-759.7 (681.9)	-513.7 (680.4)
Texas State-San Marcos	-12666.5 (247.8)	-7687.0 (254.0)	-2758.7 (556.2)	-2484.9 (549.7)	-2439.3 (549.9)	-1846.2 (546.4)
Stephen F. Austin State	-14997.2 (257.3)	-8446.1 (264.9)	-874.2 (624.7)	-915.1 (616.3)	-896.9 (616.2)	-743.8 (614.4)
Sul Ross State	-19416.6 (583.4)	-9294.4 (578.1)	1765.8 (1743.4)	1229.5 (1701.7)	1144.3 (1692.7)	1650.0 (1690.4)
Prairie View A&M	-22478.4 (307.7)	-7298.8 (335.5)	3999.5 (929.3)	3407.2 (908.0)	3334.4 (905.4)	3255.3 (904.2)
Tarleton State	-15148.5 (307.5)	-10201.1 (312.1)	-1880.6 (1029.8)	-1991.5 (1009.6)	-1987.6 (1010.0)	-1770.2 (1007.9)
TAMU	1117.2 (238.7)	1180.4 (234.8)	1142.5 (504.3)	1098.6 (495.8)	1109.3 (495.9)	1359.1 (491.2)
TAMU-Kingsville	-15918.0 (394.5)	-6733.6 (390.7)	810.1 (1017.3)	1001.0 (987.2)	1001.9 (985.0)	977.4 (984.3)
Texas Southern	-27069.8 (285.7)	-10110.6 (320.5)	1500.2 (915.2)	1621.3 (895.7)	1607.4 (894.4)	1279.0 (889.7)
Texas Tech	-6073.4 (257.1)	-3404.1 (260.9)	493.2 (611.8)	364.2 (604.0)	440.3 (604.1)	626.3 (601.5)
Texas Woman's	-15894.9 (399.1)	-5234.6 (395.5)	1654.0 (1022.4)	1464.0 (1006.5)	1458.8 (1003.6)	1488.1 (1001.7)
Houston	-10512.4 (263.3)	-4958.3 (266.5)	-626.8 (627.3)	-630.4 (618.3)	-617.4 (618.3)	-999.4 (617.1)
UT-Arlington	-12736.2 (285.2)	-8073.9 (287.8)	-1723.9 (678.7)	-1775.6 (671.5)	-1810.7 (671.2)	-1821.1 (669.5)
UT-El Paso	-22868.3 (276.2)	-12038.4 (293.4)	-706.7 (1195.9)	-158.8 (1192.5)	-199.6 (1191.7)	121.4 (1185.1)
West Texas A&M	-16246.6 (365.0)	-11177.6 (362.0)	612.8 (1427.0)	100.4 (1390.2)	104.3 (1388.6)	419.3 (1391.0)
TAMU-International	-18883.9 (422.1)	-8237.7 (423.1)	544.5 (1324.9)	417.4 (1288.0)	384.1 (1287.0)	752.1 (1284.7)
UT-Dallas	-6137.2 (429.5)	-5858.8 (426.6)	-2238.2 (1021.9)	-3426.6 (1014.0)	-3453.5 (1014.1)	-3831.6 (1016.8)
UT-Permian Basin	-12990.4 (645.6)	-6052.9 (627.8)	6408.9 (2280.6)	5898.5 (2244.6)	5938.0 (2243.6)	4906.5 (2236.0)
UT-San Antonio	-16326.4 (238.7)	-9185.6 (247.8)	-2079.0 (565.2)	-2193.8 (557.0)	-2210.7 (556.7)	-1895.2 (554.5)
TAMU-Galveston	-7705.1 (687.9)	-4774.3 (674.0)	1665.0 (1439.6)	1282.4 (1406.3)	1306.2 (1405.0)	1170.4 (1394.8)
TAMU-Corpus Christi	-15723.0 (334.3)	-8499.4 (334.3)	299.8 (770.6)	-200.0 (755.3)	-135.6 (754.7)	-26.61 (753.4)
UT-Tyler	-12125.9 (513.4)	-8843.8 (501.7)	-2680.7 (1367.7)	-3436.9 (1324.4)	-3466.9 (1323.5)	-3070.5 (1323.1)
Houston-Downtown	-20027.1 (346.8)	-6344.2 (361.1)	3043.0 (1212.2)	3808.2 (1201.1)	3723.4 (1201.4)	2874.8 (1200.2)
R-Squared	0.0777	0.1236	0.0973	0.1322	0.1335	0.1527
N	358658	358658	354403	354403	354403	354306

Notes: These estimates correspond to Figures 2 and 4. Each column presents point estimates and robust standard errors from a regression of individual student outcomes on college treatment indicators, omitting UT-Austin as the reference treatment. The UT-Austin outcome mean is 55,975. All specifications control for cohort fixed effects. The Raw Means specification controls for nothing else. The Typical Controls specification adds controls for demographics (gender, race, FRPL), high school academic preparation (10th grade test scores, advanced coursework, and top high school GPA decile indicator), and behavioral measures of non-cognitive skills (high school attendance, disciplinary infractions, and an indicator for ever being at risk of dropping out). The Baseline Specification controls solely for college admission portfolio fixed effects (and cohort fixed effects). The Main Specification adds our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. Extra Control Set 1 adds controls for advanced high school coursework, disciplinary infractions, and an indicator for ever being at risk of dropping out. Extra Control Set 2 adds fixed effects for every high school and an indicator for being in the top decile of high school GPA.

Table B.3: Balance Across Observed Ability Measures

	10th Grade Math Score		10th Grade Reading Score	
	Raw Means	Baseline Spec	Raw Means	Baseline Spec
Angelo State	-1.041 (0.0104)	-0.0102 (0.0335)	-0.848 (0.0109)	0.0273 (0.0358)
TAMU-Commerce	-0.969 (0.0156)	0.00862 (0.0425)	-0.725 (0.0156)	-0.0117 (0.0432)
Lamar	-1.143 (0.0108)	0.00709 (0.0294)	-0.875 (0.0108)	0.0469 (0.0302)
Midwestern State	-1.028 (0.0134)	-0.00710 (0.0364)	-0.767 (0.0134)	0.0623 (0.0353)
North Texas	-0.670 (0.00604)	0.000533 (0.0144)	-0.407 (0.00596)	0.0795 (0.0145)
UT-Pan American	-1.255 (0.00918)	-0.0428 (0.0274)	-1.149 (0.00997)	-0.0519 (0.0289)
Sam Houston State	-0.883 (0.00775)	-0.0764 (0.0190)	-0.647 (0.00781)	-0.0272 (0.0194)
Texas State-San Marcos	-0.643 (0.00575)	-0.122 (0.0129)	-0.400 (0.00576)	-0.0425 (0.0130)
Stephen F. Austin State	-0.946 (0.00775)	-0.0270 (0.0175)	-0.686 (0.00773)	0.000393 (0.0177)
Sul Ross State	-1.624 (0.0260)	0.00530 (0.0674)	-1.532 (0.0295)	-0.0790 (0.0804)
Prairie View A&M	-1.932 (0.0144)	0.0409 (0.0394)	-1.676 (0.0154)	0.00812 (0.0417)
Tarleton State	-0.990 (0.00992)	0.00520 (0.0303)	-0.767 (0.0104)	0.0978 (0.0317)
TAMU	-0.105 (0.00364)	-0.0628 (0.00749)	-0.0859 (0.00391)	-0.0328 (0.00838)
TAMU-Kingsville	-1.263 (0.0145)	-0.0622 (0.0325)	-1.165 (0.0154)	-0.0281 (0.0349)
Texas Southern	-2.042 (0.0137)	0.0734 (0.0369)	-1.801 (0.0153)	0.0167 (0.0391)
Texas Tech	-0.501 (0.00513)	-0.0868 (0.0113)	-0.365 (0.00519)	-0.0261 (0.0119)
Texas Woman's	-1.180 (0.0164)	-0.00885 (0.0397)	-0.761 (0.0159)	0.0433 (0.0388)
Houston	-0.570 (0.00613)	0.00637 (0.0136)	-0.555 (0.00662)	0.0150 (0.0151)
UT-Arlington	-0.602 (0.00743)	0.0228 (0.0172)	-0.527 (0.00781)	0.0271 (0.0183)
UT-El Paso	-1.220 (0.00959)	-0.0115 (0.0342)	-1.093 (0.0103)	-0.00773 (0.0352)
West Texas A&M	-0.872 (0.0122)	0.0409 (0.0384)	-0.720 (0.0130)	0.0488 (0.0404)
TAMU-International	-1.156 (0.0178)	0.0752 (0.0502)	-1.105 (0.0196)	0.00127 (0.0545)
UT-Dallas	-0.0242 (0.00767)	0.111 (0.0169)	-0.121 (0.00893)	0.0907 (0.0204)
UT-Permian Basin	-0.937 (0.0223)	-0.0172 (0.0735)	-0.769 (0.0235)	-0.0870 (0.0788)
UT-San Antonio	-0.883 (0.00627)	-0.0133 (0.0140)	-0.687 (0.00639)	0.0336 (0.0146)
TAMU-Galveston	-0.580 (0.0153)	-0.0747 (0.0303)	-0.494 (0.0159)	-0.00489 (0.0332)
TAMU-Corpus Christi	-0.909 (0.0105)	-0.00876 (0.0231)	-0.672 (0.0108)	0.0616 (0.0242)
UT-Tyler	-0.499 (0.0140)	0.0315 (0.0353)	-0.327 (0.0135)	0.0706 (0.0376)
Houston-Downtown	-1.696 (0.0162)	0.0674 (0.0489)	-1.559 (0.0180)	-0.0184 (0.0578)
R-Squared	0.2419	0.2772	0.1832	0.2105
N	422956	418267	422956	418267

Notes: These estimates correspond to Figure 3. Each column presents point estimates and robust standard errors from a regression of individual student standardized test scores on college treatment indicators, omitting UT-Austin as the reference treatment. The UT-Austin outcome mean is 0.68 SDs for Math and 0.55 SDs for Reading. All specifications control for cohort fixed effects. The Raw Means specification controls for nothing else. The Baseline Specification adds college admission portfolio fixed effects.

Table B.4: Alternative Specifications of Admissions Portfolios: BA Completion

	Main Spec	Additive	Apps Only	2+ Admits	Apps Too	Port-Top10	Port-X
Angelo State	-0.0417 (0.0175)	-0.296 (0.0117)	-0.0975 (0.0128)	-0.0431 (0.0176)	-0.0238 (0.0190)	-0.0372 (0.0177)	-0.0333 (0.0194)
TAMU-Commerce	-0.0260 (0.0225)	-0.223 (0.0141)	-0.0767 (0.0185)	-0.0250 (0.0227)	-0.0108 (0.0244)	-0.0373 (0.0232)	-0.0129 (0.0257)
Lamar	-0.0360 (0.0151)	-0.292 (0.00881)	-0.110 (0.0136)	-0.0346 (0.0152)	-0.0514 (0.0161)	-0.0427 (0.0153)	-0.0441 (0.0166)
Midwestern State	-0.0217 (0.0193)	-0.251 (0.0111)	-0.0697 (0.0171)	-0.0208 (0.0195)	-0.0122 (0.0211)	-0.0255 (0.0196)	-0.0342 (0.0210)
North Texas	-0.0317 (0.00831)	-0.190 (0.00503)	-0.0589 (0.00714)	-0.0314 (0.00838)	-0.0344 (0.00863)	-0.0322 (0.00839)	-0.0332 (0.00876)
UT-Pan American	-0.00467 (0.0143)	-0.228 (0.00905)	-0.0495 (0.0132)	-0.00597 (0.0145)	-0.00670 (0.0151)	-0.00424 (0.0144)	0.0000583 (0.0157)
Sam Houston State	0.0293 (0.0105)	-0.162 (0.00647)	-0.0124 (0.00898)	0.0286 (0.0106)	0.0249 (0.0110)	0.0274 (0.0107)	0.0280 (0.0113)
Texas State-San Marcos	-0.0308 (0.00757)	-0.158 (0.00480)	-0.0420 (0.00605)	-0.0324 (0.00765)	-0.0251 (0.00782)	-0.0315 (0.00764)	-0.0332 (0.00797)
Stephen F. Austin State	-0.00467 (0.00955)	-0.197 (0.00559)	-0.0502 (0.00815)	-0.00437 (0.00964)	-0.00736 (0.0100)	-0.0103 (0.00966)	-0.000516 (0.0101)
Sul Ross State	-0.00306 (0.0321)	-0.234 (0.0183)	-0.0539 (0.0285)	-0.00460 (0.0324)	-0.00423 (0.0349)	-0.0129 (0.0325)	0.0169 (0.0366)
Prairie View A&M	0.118 (0.0177)	-0.156 (0.0116)	0.0149 (0.0157)	0.119 (0.0178)	0.106 (0.0193)	0.106 (0.0180)	0.113 (0.0198)
Tarleton State	0.0162 (0.0170)	-0.198 (0.0112)	-0.0452 (0.0138)	0.0177 (0.0172)	0.00192 (0.0183)	0.0128 (0.0174)	0.0262 (0.0189)
TAMU	0.0267 (0.00455)	-0.0554 (0.00324)	0.0302 (0.00401)	0.0261 (0.00459)	0.0247 (0.00460)	0.0279 (0.00457)	0.0271 (0.00465)
TAMU-Kingsville	0.00696 (0.0161)	-0.207 (0.00977)	-0.0579 (0.0149)	0.00684 (0.0163)	0.00686 (0.0173)	0.00619 (0.0162)	0.00788 (0.0178)
Texas Southern	-0.0531 (0.0166)	-0.309 (0.00827)	-0.150 (0.0133)	-0.0533 (0.0167)	-0.0722 (0.0178)	-0.0583 (0.0168)	-0.0533 (0.0180)
Texas Tech	0.0131 (0.00675)	-0.104 (0.00414)	0.0128 (0.00553)	0.0118 (0.00682)	0.0161 (0.00695)	0.0103 (0.00683)	0.0138 (0.00703)
Texas Woman's	-0.0191 (0.0200)	-0.199 (0.0135)	-0.0734 (0.0165)	-0.0182 (0.0202)	-0.0138 (0.0219)	-0.0176 (0.0205)	-0.00771 (0.0229)
Houston	-0.0835 (0.00831)	-0.242 (0.00539)	-0.109 (0.00719)	-0.0838 (0.00838)	-0.0800 (0.00863)	-0.0852 (0.00838)	-0.0830 (0.00872)
UT-Arlington	-0.0525 (0.0104)	-0.216 (0.00707)	-0.0911 (0.00911)	-0.0521 (0.0105)	-0.0680 (0.0110)	-0.0522 (0.0105)	-0.0495 (0.0112)
UT-El Paso	-0.0685 (0.0186)	-0.263 (0.0111)	-0.115 (0.0170)	-0.0702 (0.0188)	-0.0699 (0.0196)	-0.0647 (0.0187)	-0.0550 (0.0205)
West Texas A&M	-0.0478 (0.0216)	-0.258 (0.0134)	-0.105 (0.0184)	-0.0467 (0.0218)	-0.0366 (0.0228)	-0.0471 (0.0220)	-0.0478 (0.0238)
TAMU-International	0.0616 (0.0240)	-0.165 (0.0160)	0.00567 (0.0220)	0.0629 (0.0243)	0.0372 (0.0257)	0.0577 (0.0246)	0.0533 (0.0274)
UT-Dallas	-0.00157 (0.0122)	-0.117 (0.00927)	-0.00611 (0.0103)	0.000855 (0.0123)	-0.00103 (0.0126)	-0.00357 (0.0125)	0.00894 (0.0130)
UT-Permian Basin	0.00317 (0.0382)	-0.237 (0.0221)	-0.0555 (0.0348)	0.00205 (0.0386)	0.0153 (0.0419)	0.00355 (0.0396)	0.0467 (0.0451)
UT-San Antonio	-0.0614 (0.00806)	-0.239 (0.00513)	-0.114 (0.00674)	-0.0613 (0.00814)	-0.0762 (0.00864)	-0.0596 (0.00814)	-0.0551 (0.00856)
TAMU-Galveston	0.0301 (0.0191)	-0.127 (0.0132)	-0.0218 (0.0158)	0.0297 (0.0193)	0.0345 (0.0203)	0.0237 (0.0195)	0.0315 (0.0205)
TAMU-Corpus Christi	-0.0236 (0.0127)	-0.195 (0.00808)	-0.0535 (0.0113)	-0.0231 (0.0128)	-0.0290 (0.0135)	-0.0221 (0.0128)	-0.0111 (0.0138)
UT-Tyler	-0.0386 (0.0224)	-0.216 (0.0147)	-0.0800 (0.0206)	-0.0371 (0.0227)	-0.0649 (0.0247)	-0.0489 (0.0230)	-0.0513 (0.0258)
Houston-Downtown	-0.0431 (0.0247)	-0.322 (0.0164)	-0.185 (0.0196)	-0.0446 (0.0249)	-0.0496 (0.0270)	-0.0388 (0.0248)	-0.0336 (0.0278)
R-Squared	0.2166	0.2058	0.2213	0.1889	0.2287	0.2295	0.2423
N	418267	422956	415668	125876	408609	418267	418267

Notes: These estimates correspond to Figures A.2 and A.3. All specifications control for cohort fixed effects. The Main Specification controls for admission portfolio fixed effects and our core set of covariates: gender, FRPL, race, 10th grade test scores, and high school attendance. Additive replaces the full set of college admission portfolio fixed effects from the Main Specification with 29 dummies indicating an application to each college, and 29 dummies indicating an admission to each college, with UT-Austin as the omitted category. Apps Only replaces the college admission portfolio fixed effects from the Main Specification with fixed effects solely for each combination of applications, ignoring information on admissions. 2+ Admits runs the Main Specification on the subsample of students who are admitted to at least two colleges. Apps Too replaces the admission portfolio fixed effects with fixed effects for every distinct portfolio of applications and admissions. Port-Top10 interacts the admission portfolio indicators with the indicator for being in top decile of high school GPA. Port-X interacts the admission portfolio indicators with our core covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance.

Table B.5: Alternative Specifications of Admissions Portfolios: Earnings

	Main Spec	Additive	Apps Only	2+ Admits	Apps Too	Port-Top10	Port-X
Angelo State	-1716.1 (1056.2)	-12057.8 (685.3)	-4457.3 (790.0)	-1639.8 (1070.0)	-1334.8 (1144.7)	-1498.1 (1081.0)	-2258.0 (1193.4)
TAMU-Commerce	-2854.1 (1160.7)	-13155.0 (701.1)	-4939.7 (943.3)	-2925.0 (1172.6)	-2259.9 (1281.9)	-3359.1 (1188.9)	-2005.5 (1327.4)
Lamar	311.4 (880.8)	-9590.9 (505.0)	-2856.8 (773.6)	309.6 (890.8)	362.7 (963.7)	432.4 (896.6)	407.1 (975.1)
Midwestern State	-1789.6 (1072.3)	-13050.9 (598.8)	-4907.6 (917.9)	-1814.0 (1083.3)	-2734.3 (1178.0)	-1607.3 (1098.5)	-2266.3 (1187.9)
North Texas	-1467.1 (568.0)	-9048.5 (319.4)	-3523.9 (466.2)	-1485.6 (574.6)	-1691.5 (587.6)	-1507.7 (575.6)	-1490.3 (602.1)
UT-Pan American	895.5 (871.1)	-10016.4 (505.4)	-1354.3 (773.0)	891.2 (880.8)	1422.3 (927.7)	1164.8 (884.1)	1419.7 (935.6)
Sam Houston State	-739.6 (682.4)	-8907.0 (391.2)	-1959.1 (555.2)	-818.8 (690.1)	-683.2 (712.8)	-721.8 (690.7)	-731.1 (730.7)
Texas State-San Marcos	-2484.9 (549.7)	-8512.7 (335.4)	-3554.9 (429.1)	-2494.5 (556.5)	-2417.2 (570.5)	-2403.8 (559.0)	-2228.2 (581.8)
Stephen F. Austin State	-915.1 (616.3)	-9632.4 (342.1)	-3286.1 (508.5)	-990.1 (623.4)	-1366.6 (652.2)	-1062.1 (626.3)	-805.9 (660.7)
Sul Ross State	1229.5 (1701.7)	-7341.0 (975.5)	-1511.0 (1473.5)	1270.5 (1721.2)	1651.5 (1878.7)	1019.7 (1731.8)	2893.5 (1969.3)
Prairie View A&M	3407.2 (908.0)	-8956.2 (549.3)	-1399.2 (782.9)	3349.7 (918.0)	3679.3 (1002.5)	2772.7 (908.5)	3606.2 (987.1)
Tarleton State	-1991.5 (1009.6)	-10200.7 (676.9)	-4375.4 (818.4)	-2067.1 (1020.7)	-2473.7 (1087.9)	-2206.8 (1033.1)	-2112.9 (1115.8)
TAMU	1098.6 (495.8)	-4324.7 (337.3)	832.1 (407.5)	1076.0 (501.5)	963.8 (507.6)	1066.0 (502.5)	1037.9 (521.3)
TAMU-Kingsville	1001.0 (987.2)	-8696.8 (589.1)	-1638.1 (921.3)	953.3 (998.7)	933.4 (1037.1)	998.0 (995.9)	683.5 (1056.8)
Texas Southern	1621.3 (895.7)	-11305.7 (420.9)	-3796.3 (691.8)	1563.9 (905.7)	1778.3 (980.4)	1570.6 (908.4)	1959.1 (962.1)
Texas Tech	364.2 (604.0)	-6809.1 (344.7)	-1054.4 (471.0)	334.6 (610.8)	372.4 (631.1)	326.6 (611.7)	217.3 (639.5)
Texas Woman's	1464.0 (1006.5)	-8127.0 (661.0)	-1861.6 (816.7)	1373.8 (1016.6)	1176.3 (1099.8)	1697.9 (1030.2)	1868.5 (1103.1)
Houston	-630.4 (618.3)	-7603.0 (372.7)	-2589.7 (513.7)	-652.3 (625.1)	-895.1 (646.6)	-747.4 (627.2)	-627.4 (661.9)
UT-Arlington	-1775.6 (671.5)	-9124.9 (441.6)	-4242.5 (580.2)	-1797.1 (678.5)	-2395.4 (710.2)	-1882.6 (680.2)	-1789.9 (729.8)
UT-El Paso	-158.8 (1192.5)	-11537.2 (720.9)	-3613.7 (1074.7)	-123.3 (1204.4)	70.87 (1283.0)	-327.9 (1195.8)	731.9 (1264.2)
West Texas A&M	100.4 (1390.2)	-10345.2 (801.1)	-2690.7 (1169.1)	82.18 (1406.8)	-71.91 (1499.0)	689.6 (1402.4)	-299.9 (1535.9)
TAMU-International	417.4 (1288.0)	-8876.8 (880.9)	-1611.0 (1185.3)	398.2 (1300.6)	-397.9 (1414.2)	393.6 (1303.8)	190.6 (1349.3)
UT-Dallas	-3426.6 (1014.0)	-7927.1 (723.5)	-3698.5 (808.5)	-3434.8 (1024.6)	-3960.1 (1058.3)	-3526.8 (1032.2)	-3315.5 (1125.0)
UT-Permian Basin	5898.5 (2244.6)	-5693.8 (1190.6)	267.5 (2074.9)	5903.1 (2271.5)	5067.6 (2400.8)	7510.0 (2279.8)	9679.7 (2525.7)
UT-San Antonio	-2193.8 (557.0)	-10095.9 (326.2)	-4387.3 (457.1)	-2220.9 (563.2)	-2514.1 (586.9)	-2167.2 (566.4)	-1852.7 (599.2)
TAMU-Galveston	1282.4 (1406.3)	-4646.9 (1019.4)	-812.5 (1197.7)	1253.2 (1422.3)	1271.1 (1554.3)	1338.0 (1420.2)	833.8 (1584.1)
TAMU-Corpus Christi	-200.0 (755.3)	-9011.0 (459.3)	-2171.2 (671.9)	-243.5 (764.4)	1.338 (807.9)	-166.6 (764.9)	57.73 (822.0)
UT-Tyler	-3436.9 (1324.4)	-11005.7 (871.7)	-5032.4 (1212.2)	-3459.8 (1340.3)	-4425.0 (1455.2)	-3609.7 (1364.8)	-2967.8 (1524.5)
Houston-Downtown	3808.2 (1201.1)	-6446.2 (841.3)	-3612.9 (1005.9)	3932.9 (1215.3)	3429.5 (1344.0)	3911.0 (1211.8)	3786.0 (1330.3)
R-Squared	0.1322	0.1219	0.1346	0.1401	0.1408	0.1404	0.1611
N	354403	358658	352029	106857	345621	354403	354403

Notes: These estimates correspond to Figures A.2 and A.3. All specifications control for cohort fixed effects. The Main Specification controls for admission portfolio fixed effects and our core set of covariates: gender, FRPL, race, 10th grade test scores, and high school attendance. Additive replaces the full set of college admission portfolio fixed effects from the Main Specification with 29 dummies indicating an application to each college, and 29 dummies indicating an admission to each college, with UT-Austin as the omitted category. Apps Only replaces the college admission portfolio fixed effects from the Main Specification with fixed effects solely for each combination of applications, ignoring information on admissions. 2+ Admits runs the Main Specification on the subsample of students who are admitted to at least two colleges. Apps Too replaces the admission portfolio fixed effects with fixed effects for every distinct portfolio of applications and admissions. Port-Top10 interacts the admission portfolio indicators with the indicator for being in top decile of high school GPA. Port-X interacts the admission portfolio indicators with our core covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance.

Table B.6: Student-Level Regression Estimates of the Return to Attending a More Selective College

	Raw	Typical	Baseline	Main Spec	Extra 1	Extra 2	UTA Admits
College Mean SAT Score (x100)	6839.9 (45.79)	3141.0 (55.50)	3.326 (154.4)	-27.69 (152.0)	-24.68 (151.9)	-72.78 (151.4)	-103.9 (269.3)
R-Squared	0.0670	0.1149	0.0970	0.1319	0.1331	0.1524	0.0717
N	358658	358658	354403	354403	354403	354306	52,459

Notes: These estimates correspond to Figure 6. The mean of the dependent variable (annualized earnings) is \$44,834. Each point estimate and robust 95 percent confidence interval comes from a regression of individual student earnings on the mean incoming SAT score of the student's college. The coefficients are scaled to correspond to a 100-point increase in mean SAT scores. All specifications control for cohort fixed effects. The Raw specification controls for nothing else. The Typical Controls specification adds controls for demographics (gender, race, FRPL), high school academic preparation (10th grade test scores, advanced coursework, and top high school GPA decile indicator), and behavioral measures of non-cognitive skills (high school attendance, disciplinary infractions, and an indicator for ever being at risk of dropping out). The Baseline Specification controls solely for college admission portfolio fixed effects (and cohort fixed effects). The Main Specification adds our core set of covariate controls: gender, FRPL, race, 10th grade test scores, and high school attendance. Extra Control Set 1 adds controls for advanced high school coursework, disciplinary infractions, and an indicator for ever being at risk of dropping out. Extra Control Set 2 adds fixed effects for every high school and an indicator for being in the top decile of high school GPA. The final estimate comes from running the Main Specification on the subsample of students who receive admission to UT-Austin.

Table B.7: Match Effects Specification with Treatment-Covariate Interactions: BA Completion

	Main Spec	Add Portfolio-Covariate Interactions	Add Treatment-Covariate Interactions
Angelo State	-0.0417 (0.0175)	-0.0333 (0.0194)	-0.0468 (0.0214)
TAMU-Commerce	-0.0260 (0.0225)	-0.0129 (0.0257)	0.00359 (0.0323)
Lamar	-0.0360 (0.0151)	-0.0441 (0.0166)	-0.0254 (0.0227)
Midwestern State	-0.0217 (0.0193)	-0.0342 (0.0210)	-0.0473 (0.0243)
North Texas	-0.0317 (0.00831)	-0.0332 (0.00876)	-0.0310 (0.0115)
UT-Pan American	-0.00467 (0.0143)	0.0000583 (0.0157)	-0.00619 (0.0413)
Sam Houston State	0.0293 (0.0105)	0.0280 (0.0113)	0.0265 (0.0135)
Texas State-San Marcos	-0.0308 (0.00757)	-0.0332 (0.00797)	-0.0328 (0.0107)
Stephen F. Austin State	-0.00467 (0.00955)	-0.000516 (0.0101)	-0.00709 (0.0133)
Sul Ross State	-0.00306 (0.0321)	0.0169 (0.0366)	0.0358 (0.0488)
Prairie View A&M	0.118 (0.0177)	0.113 (0.0198)	0.315 (0.111)
Tarleton State	0.0162 (0.0170)	0.0262 (0.0189)	0.00686 (0.0282)
TAMU	0.0267 (0.00455)	0.0271 (0.00465)	0.0276 (0.00898)
TAMU-Kingsville	0.00696 (0.0161)	0.00788 (0.0178)	-0.0449 (0.0260)
Texas Southern	-0.0531 (0.0166)	-0.0533 (0.0180)	-0.0989 (0.0891)
Texas Tech	0.0131 (0.00675)	0.0138 (0.00703)	0.0126 (0.0110)
Texas Woman's	-0.0191 (0.0200)	-0.00771 (0.0229)	-0.0278 (0.0553)
Houston	-0.0835 (0.00831)	-0.0830 (0.00872)	-0.0934 (0.0111)
UT-Arlington	-0.0525 (0.0104)	-0.0495 (0.0112)	-0.0455 (0.0133)
UT-El Paso	-0.0685 (0.0186)	-0.0550 (0.0205)	-0.0930 (0.0318)
West Texas A&M	-0.0478 (0.0216)	-0.0478 (0.0238)	-0.0409 (0.0311)
TAMU-International	0.0616 (0.0240)	0.0533 (0.0274)	0.311 (0.102)
UT-Dallas	-0.00157 (0.0122)	0.00894 (0.0130)	0.0238 (0.0221)
UT-Permian Basin	0.00317 (0.0382)	0.0467 (0.0451)	0.109 (0.0518)
UT-San Antonio	-0.0614 (0.00806)	-0.0551 (0.00856)	-0.0595 (0.0107)
TAMU-Galveston	0.0301 (0.0191)	0.0315 (0.0205)	0.0137 (0.0352)
TAMU-Corpus Christi	-0.0236 (0.0127)	-0.0111 (0.0138)	-0.00701 (0.0169)
UT-Tyler	-0.0386 (0.0224)	-0.0513 (0.0258)	-0.0725 (0.0336)
Houston-Downtown	-0.0431 (0.0247)	-0.0336 (0.0278)	-0.0724 (0.0459)
R-Squared	0.2166	0.2423	0.2428
N	418267	418267	418267

Notes: These estimates correspond to Figure A.5. Each column presents point estimates and robust standard errors from a regression of individual student outcomes on college treatment indicators, omitting UT-Austin as the reference treatment. The UT-Austin outcome mean is 0.82. All specifications control for cohort fixed effects. The Main Specification controls for admission portfolio fixed effects and our core set of covariates: gender, FRPL, race, 10th grade test scores, and high school attendance. The Portfolio-Covariate Interactions specification adds interactions between the admission portfolios and the core covariates. The Treatment-Covariate Interactions specification further adds interactions between the college treatments and the core covariates.

Table B.8: Match Effects Specification with Treatment-Covariate Interactions: Earnings

	Main Spec	Add Portfolio-Covariate Interactions	Add Treatment-Covariate Interactions
Angelo State	-1716.1 (1056.2)	-2258.0 (1193.4)	-2918.0 (1320.2)
TAMU-Commerce	-2854.1 (1160.7)	-2005.5 (1327.4)	-1858.2 (1633.1)
Lamar	311.4 (880.8)	407.1 (975.1)	-770.1 (1419.3)
Midwestern State	-1789.6 (1072.3)	-2266.3 (1187.9)	-1907.5 (1342.4)
North Texas	-1467.1 (568.0)	-1490.3 (602.1)	-1393.6 (705.4)
UT-Pan American	895.5 (871.1)	1419.7 (935.6)	4949.3 (2991.4)
Sam Houston State	-739.6 (682.4)	-731.1 (730.7)	-681.1 (855.2)
Texas State-San Marcos	-2484.9 (549.7)	-2228.2 (581.8)	-2310.6 (679.6)
Stephen F. Austin State	-915.1 (616.3)	-805.9 (660.7)	-1042.0 (799.9)
Sul Ross State	1229.5 (1701.7)	2893.5 (1969.3)	-453.5 (2862.9)
Prairie View A&M	3407.2 (908.0)	3606.2 (987.1)	23173.9 (6872.6)
Tarleton State	-1991.5 (1009.6)	-2112.9 (1115.8)	-1720.8 (1494.1)
TAMU	1098.6 (495.8)	1037.9 (521.3)	1111.0 (682.3)
TAMU-Kingsville	1001.0 (987.2)	683.5 (1056.8)	-503.4 (1746.1)
Texas Southern	1621.3 (895.7)	1959.1 (962.1)	17295.4 (6288.1)
Texas Tech	364.2 (604.0)	217.3 (639.5)	47.23 (777.8)
Texas Woman's	1464.0 (1006.5)	1868.5 (1103.1)	-496.2 (3125.6)
Houston	-630.4 (618.3)	-627.4 (661.9)	-292.4 (753.3)
UT-Arlington	-1775.6 (671.5)	-1789.9 (729.8)	-1259.1 (794.6)
UT-El Paso	-158.8 (1192.5)	731.9 (1264.2)	1034.4 (2320.6)
West Texas A&M	100.4 (1390.2)	-299.9 (1535.9)	-239.0 (1855.3)
TAMU-International	417.4 (1288.0)	190.6 (1349.3)	-3754.7 (5927.1)
UT-Dallas	-3426.6 (1014.0)	-3315.5 (1125.0)	-1795.0 (1547.6)
UT-Permian Basin	5898.5 (2244.6)	9679.7 (2525.7)	12695.0 (3199.7)
UT-San Antonio	-2193.8 (557.0)	-1852.7 (599.2)	-1669.8 (707.3)
TAMU-Galveston	1282.4 (1406.3)	833.8 (1584.1)	-4795.6 (2024.3)
TAMU-Corpus Christi	-200.0 (755.3)	57.73 (822.0)	-796.1 (1015.6)
UT-Tyler	-3436.9 (1324.4)	-2967.8 (1524.5)	-3605.5 (1896.9)
Houston-Downtown	3808.2 (1201.1)	3786.0 (1330.3)	2483.2 (2277.1)
R-Squared	0.1322	0.1611	0.1616
N	354403	354403	354403

Notes: These estimates correspond to Figure A.5. Each column presents point estimates and robust standard errors from a regression of individual student outcomes on college treatment indicators, omitting UT-Austin as the reference treatment. The UT-Austin outcome mean is 55,975. All specifications control for cohort fixed effects. The Main Specification controls for admission portfolio fixed effects and our core set of covariates: gender, FRPL, race, 10th grade test scores, and high school attendance. The Portfolio-Covariate Interactions specification adds interactions between the admission portfolios and the core covariates. The Treatment-Covariate Interactions specification further adds interactions between the college treatments and the core covariates.