Does Higher Instructional Spending in Colleges Promote Student Earnings?

Quanfeng Zhou *
University of California, Riverside

October 18, 2021

[Latest Version]

Abstract

Are large increases in college spending worthwhile? This paper focuses on the relationship between instructional spending in colleges and student earnings after entering the labor market. By exploiting an institution-level panel data set, I use two-way fixed effects regressions to estimate the elasticities between instructional spending and student earnings. The estimated elasticities are 0.5, 1.7, and 1.9 percent for average earnings six, eight, and ten years after first attendance. They are also higher on lower percentiles in the earning distribution. Compared to the estimates in the existing literature, my estimates are much smaller. I further interact instructional spending with the type of institution. I find that the results are driven mainly through private institutions and four-year institutions. Public institutions that are more established are also found to have a positive, but smaller effect. Cost-benefit analysis reveals that instructional spending is cost-effective only under an extremely forward-looking perspective.

Keywords: instructional spending, college quality, student earnings

*Department of Economics, University of California Riverside (email: quanfeng.zhou@email.ucr.edu)

1 Introduction

More money is spent on fewer college students in the United States now than a decade ago. In real terms, between academic year (AY) 2009-10 and 2018-19, expenditure per full-time equivalent (FTE) student increased by 24.6% for public institutions and 16.4% for private nonprofit institutions.[1] The large increases in per-student spending result from both the growth of total spending and the decline in enrollment. In AY2018-19, all postsecondary institutions spent a total of $626 billion on instruction, institutional support, research, academic support, and other services, rising from $535 billion in AY2009-2010. Over the same period, fall enrollment for degree-seeking students declined from 21 million to 19.6 million.[2] With the large increases in per-student spending, it is natural to ask whether these increases are worthwhile.

Instructional spending is the largest spending category for most institutions, while it has been decreasing as a share of total spending.[3] Figure 1 depicts the growth of instructional spending and total spending. We see that its growth lagged the growth of total spending, and indeed, its share decreased from 41% in 1986 to 29% in 2018. What might justify that institutions are allocating relatively fewer funds to instruction? This paper investigates the (in)effectiveness of instructional in promoting students' earnings after graduation.

Figure 2 plots the percentage change in average wage earnings for young workers (age 22-30) with at least associate degree between 2012 and 2019 against the percentage change in instructional spending per college student between 2009 and 2016 in each state. The dashed line is a linear fit weighted by the population of such young workers in 2019. We

[1] Table 334.10 and Table 334.30, Digest of Education Statistics 2020, National Center for Education Statistic (NCES). In 2019 constant dollar, expenditure per full-time equivalent (FTE) student was $31,076 in AY2009-10, and $38,709 in AY2018-19 for public institutions. In private institutions, expenditure per FTE student was $54,416 in AY2009-10, and $63,321 in AY2018-19.

[2] Total spending is in 2019 constant dollars. Author's calculation based on data from the Integrated Postsecondary Education Data System (IPEDS). Fall enrollments are retrieved from Table 303.10, Digest of Education Statistics 2020, National Center for Education Statistic.

[3] According to the IPEDS, instructional spending consists of expenses related to general academic instruction, occupational and vocational instruction, community education, preparatory and adult basic education, and regular, special, and extension sessions. It may also consist of research, public service, and information technology related expenses if an institution does not separately budget those expenses.

observe a clear positive correlation between the larger increase in spending and the larger increase in wage earnings of young workers.[4] The question is whether such a positive association is caused by any confounding factors.

To address the above issue, this paper employs the two-way fixed effects regression using data from the College Scorecard project to estimate the impact of instructional spending on student earnings. In the main specification, I include the institution fixed effects, cohort fixed effects, state-specific time trends, and a rich set of control variables. In addition to the average earnings, I also examine the effects over the 25th, 50th, and 75th percentiles in the earning distribution, for earnings six, eight, and ten years after their first attendance. Therefore, this paper draws a complete picture of the relationship between instructional spending and student earnings in the short term after graduation.

Under the main specification, I find that the elasticities of instructional spending on earning outcomes are 0.5, 1.7, and 1.9 percent for average earnings six, eight, and ten years after first attendance. They are higher on lower percentiles in the earning distribution, ranging from 0.3 to 3 percent. To put the above numbers into perspective, evaluated at the mean values of spending and earnings, a 10% increase of $806 in instructional spending (per year) will lead to an increase in wage earning by $71 eight years after a student's first attendance. After interacting instructional spending with the type of institution, I find that the results are driven mainly through private institutions and four-year institutions. The estimated elasticities are from 0.6 to 3.7 percent for private institutions, while only between 0.3 and 0.8 percent for public institutions. The estimated elasticities are between 0.6 to 5.9 percent for four-year institutions. The data that I use consist of more than 6,700 institutions, representing nearly 95% of all postsecondary institutions in the United States. Therefore, the above results are very likely to be externally valid.

To provide evidence that the estimates are not driven by confounding factors, I employ the sensitivity analysis proposed in Oster (2019) that examines to what extent the estimates can survive a confounding factor that will bring them towards zero. The results suggest

[4]The weighted correlation coefficient is 0.39. This is largely driven by the five largest states: California, Texas, Florida, New York, and Pennsylvania that are almost perfectly correlated with a correlation coefficient above 0.96, either weighted or unweighted. The five largest states also correspond to a much steeper slope. The unweighted correlation coefficient is 0.13 among all states.

that most of my estimates can survive a confounding factor of 0.5, that is, the selection on unobservables being half as strong as the selection on all observed control variables. A third of my estimates can survive a confounding factor close to 1, a level that 55% of published work using observational data cannot, as studied in Oster (2019).

This paper also uses the instrumental variable (IV) regression to estimate the effect of instructional spending on student earnings. The instrument variable, state-level budget shock multiplied by an institution's historical financial dependence on state appropriations, was constructed in Deming and Walters (2017).[5] I use the public institutions from my analysis sample for this exercise as the instrument is only applicable to public institutions. Although the instrument has high predictive power in the first stage of my treatment variable, instructional spending, the second stage was imprecisely estimated with large standard errors. The result from the IV regression is unsurprising, given that I find small and statistically insignificant effects for public institutions in the two-way fixed effects regressions.

I evaluate the cost-effectiveness of instructional spending by comparing the investment with the discounted series of earning increases, assuming my estimated effects are persistent throughout a student's lifetime. Using a highly forward looking discount factor 0.9975, a student has to work for 42 years (37 years for private institution and 58 for public institution attendees) until the net present return becomes positive. Alternatively, assuming a student works for 45 years after graduation, for the net present return to be non-negative, the discount factor needs to be 0.9946 (0.9887 for private institution and 1.0066 for public institution attendees). Commonly, we think people and the society as whole are not as forward-looking. Therefore, my results provide a possible explanation why institutions have been adjusting and lowering the share of spending on instruction.

This paper fits in the extensive literature of estimating the return to school quality. The largest strand of the literature focuses on the return to school spending before college on various outcomes, including test pass rates, college enrollment, years of education, and earnings in the labor market (Grogger (1996); Papke (2005); Lafortune et al. (2018); Hyman

[5]In their paper, they use two instrument variables (1) state-level legislative cap on tuition fee increases, and (2) budget shock multiplied by historical dependence to estimate the impact of two treatment variables (a) tuition fee, and (b) spending on college completion rate.

(2017); Jackson et al. (2016)), and conclude positive effects, but only recently. Earlier studies found that throwing money at public schools is ineffective in promoting students' performance (e.g., Hanushek (1981)). Another strand of the literature focuses on the earning returns to college quality, examining other dimensions of quality such as peer quality (Dale and Krueger (2002, 2011)) and selectivity of the college (Behrman et al. (1996); Hoekstra (2009)). A set of recent papers estimate the impact of college spending on outcomes such as enrollment, persistence, and completion (Webber and Ehrenberg (2010); Webber (2012); Deming and Walters (2017)).

This paper fills the gap by examining the effect of college spending on students' labor market earnings. We care about other outcomes, including persistence and completion, because we believe they map into earnings. There is an early wave of papers that directly estimate the impact of college spending on earnings, using cross-sectional data from Project Talent that surveys World War II veterans (Solomon (1975); Wachtel (1976); Foster and Rodgers (1980)). They estimate an elasticity coefficient between 10 to 20 percent, which is much higher than I find.

This paper is more closely related to the study by Griffith and Rask (2016). By matching individual-level data to institution characteristics from IPEDS and creating a cross-sectional dataset, they estimate the relationship between college spending and student earnings. The individual-level data are a subset of the students from the Education Longitudinal Study of 2002 (ELS:02) who had valid labor market outcomes and chose to go to four-year colleges. They employ an ordinary least square (OLS) estimation strategy and include a large set of control variables for their identification. An important variable, the average spending at other institutions the student applied to, is argued to capture the unobserved quality of the student and would hence take care of the selection of students into institutions. They use the Heckman selection model to account for

This paper exploits a newly available dataset that allows me to apply a different empirical methodology for identification. Using institution-level panel data, I include institution fixed effects that account for any time-invariant unobserved variables of institutions that could be a confounding factor and unaccounted for when using cross-sectional data. In addition, the subset of the survey participants in the ELS:02 who chose to go to college might not be nationally representative, causing concerns to both the internal and external validity of the results. The data used in this paper cover almost all major postsecondary degree-granting institutions in the United States and is much less concerned with the issue of external validity. I also include cohort fixed effects and state-specific time trends to capture general movements in labor market conditions both nationally and statewide. My identifying variation is exogenous to those confounding factors.

My results also echo a more recent paper by Mountjoy and Hickman (2021). Using more granular data from the state of Texas, they identify that increase by a standard deviation in a composite index of non-peer inputs, which includes instructional expenditures, full-time faculty share, tenure-track faculty share, and faculty-student ratio, predicts a \$753 additional earnings value-added. My results suggest that a standard deviation ($\$5,935$) increase in instructional spending (on the average of $\$8,066$) leads to an increase in average earnings by \$601, eight years after first attendance.

This paper proceeds as follows. I describe the data and background information in Section 2. Section 3 contains the empirical specification and discusses the identification strategy. The results are presented in Section 4. Section 5 considers using an instrumental variable regression approach and discusses its results. Section 6 conducts a cost-benefit analysis of the results and concludes.

2 Data and Background

I need a dataset containing both students earning outcomes and college education information to answer my research question. Ideally, linking education records to earnings nationwide generates a most comprehensive dataset for the analysis. However, such a project requires enormous resources and has not been accomplished so far to the best of my knowledge.[6] As the independent variable of interest, instructional spending per student,

[6]Many researchers studying closely related questions collaborated with local governments and universities to match the state Unemployment Insurance (UI) data to the education records. For example, Mountjoy and Hickman (2021) focus on the state of Texas. However, States vary significantly in observed and unobserved ways, raising the issue of external validity. In addition, State UI data cannot capture students who work in a state different from where they attend college.

varies only at the institution level, unlike some other variables of interest in related research questions such as student loan status and academic performance that change at the student level, this paper utilizes a nationally representative dataset from the College Scorecard project, where instructional spending per student and earning outcomes are available at the institution level for different cohorts.

The project was created in 2013 under President Barack Obama's administration to make data about colleges more accessible to the general public. It combines data from multiple existing systems and government agencies, including Integrated Postsecondary Education Data System (IPEDS), Department of Treasury, and National Student Loan Data System (NSLDS).[7] The project linked the student-level earnings data from administrative tax records maintained by the Department of the Treasury to the records in NSLDS. The linked records were then aggregated to the institution level and matched to institution characteristics that have been collected annually by the IPEDS.

There are two important advantages of using this data. First, the data are highly representative of institutions in the United States. All postsecondary institutions participating in any federal financial assistance program authorized by Title IV of the Higher Education Act are mandatory to report their data to the IPEDS, making it covering almost all postsecondary institutions in the United States. I observe at least 6,400 unique institution identifiers each year. Among all institutions, 4,355 appear in all years, of which 3,824 have valid spending information in all years, representing between 84% to 90% of the total student enrollment. My analysis sample represents 89% of all undergraduate student enrollment.[8] Second, this is an institution-level panel dataset starting from 1996 covering topics including cost, enrollment, completion, finance, demographic and socioeconomic backgrounds of students, program offerings, awards and scholarships, and many others, allowing me to control for a rich set of the institution and student characteristics.

This data also have the following complications. First, the earning outcomes are estimated for undergraduate Title IV recipients working and not enrolled in graduate programs

[7]Other agencies include Department of Education, Office of Postsecondary Education (OPE) and Federal Student Aid (FSA). They facilitated the linking process for the College Scorecard project and provided other variables that are not used in this paper.

[8]I discuss in more detail how I construct my analysis sample in Section 2.2.

when earnings are measured. The share of each institution's entering class represented by Title IV students can vary substantially due to differences in family income of students attending those institutions and state and institutional aid policies. Second, the earning statistics are calculated based on pooled two-year cohorts. On the one hand, this reduces measurement error and can lead to more stable estimates. On the other hand, it shortens the panel, making it harder to identify an effect as there are fewer variations in the independent variable. Third, earning outcomes are suppressed for confidentiality reasons for cohorts less than 30, leading to missing values in the earning outcomes for about one-fifth of the institution-cohort observations. Those are typically small institutions. Last, spending is reported based on fiscal years (Oct 1st to Sept 30th), while expenditures commonly affect students throughout academic years (Sept 1st to Aug 31st). As a check for robustness, I adjust for this difference in calculating the average spending. I complement the College Scorecard project data with the State Higher Education Finance (SHEF) project for my instrumental variable analysis.[9] The variables I use are state-level appropriations for public institutions and the number of FTE enrolled students each year. In addition, I use the variables state appropriations to each public institution and the total revenues from the finance component of IPEDS in 1990.

I use the Consumer Price Index (CPI) retrieved from the Bureau of Labor Statistics to deflate the earning variables and the Higher Education Price Index (HEPI) retrieved from Commonfund Institute to deflate the spending variables.[10] All monetary terms are in 2014 dollars unless otherwise stated.

2.1 Data Structure

The original data are a panel of institutions over the years. I transformed it into an institution-cohort panel. For example, in the 2009 data file, the variable average earning ten years after entry is essentially the average earning of the 1999 cohort, measured in 2009, ten years after they entered college. Assuming it is a four-year institution, I can go

[9]The State Higher Education Finance (SHEF) project is under the State Higher Education Executive Officers (SHEEO) Association. Deming and Walters (2017) used the same data.

[10]The HEPI is an inflation index specifically designed to reflect price levels for higher education inputs. All results are qualitatively the same if I use CPI to adjust for price level changes for my spending variable.

to the data file from 1999 to 2002 to match the spending information and other institution characteristics to the earning variables of this cohort. In addition, in the 2007 data file, the variable average earning eight years after entry measures the earnings of this same 1999 cohort in 2007.[11] Therefore, I observe multiple earning outcomes for a given cohort. Following the above transformation, I end up with a panel where earnings six, eight, and ten years after entry are measured for cohorts from 1996 to 2007 (with gaps). In addition to the average earnings, I also observe the 25th, 50th, and 75th percentile of the earnings for each cohort. Figure 3 shows the availability of the earning variables by cohort. I have a panel of eight, seven, and six cohorts for earning outcomes six, eight, and ten years after entry, respectively.

2.2 Sample Restrictions

After the transformation, the panel consists of 51,779 institution-cohort observations where the variable instructional spending and at least one of the earning outcome variables are non-missing. I call this the "Full Sample".

I focus on the institutions that remain consistent (two-year vs. four-year) levels throughout the period to avoid dramatic institutional upheavals that potentially affect spending and student earnings. Change in institution levels also leads to ambiguity in calculating the average instructional spending for a cohort. This restriction leaves out 5,125 institution-cohort observations.

Further, I exclude observations where the spending variable takes on extremely large or small values. For instance, the highest value in the data is 29 million dollars per student per year in instructional spending, which comes from a nursing school in Rhode Island. Therefore, I follow the literature and take the threshold used in Deming and Walters (2017) to admit only observations with spending between \$50 and \$100,000 per year. This restriction takes out another 4,410 institution-cohort observations, many of which are highly specialized, such as training institutions for aviation and performing arts.[12]

[11] The two cohorts may not be identical, though, as the earnings are only measured for those who are working. The employment status could differ between 2007 and 2009. The two cohorts would still overlap significantly.

[12] Their students most likely work in highly isolated labor markets. Therefore, excluding those institutions

The remaining data with 42,244 institution-cohort observations form my analysis sample, covering a total of 5,360 institutions.[13] I take the average of the instructional spending over the two years (four years) starting from the year of entry of the cohort in two-year (four-year) institutions.[14] Lastly, I average over two adjacent cohorts to correspond to the earning outcomes that are measured for two-year pooled cohorts.

2.3 Summary Statistics

Table 1 presents summary statistics for the outcome variables. In calculating the summary statistic, I weighed the sample by the undergraduate enrollment. In the full sample, the average earnings for students six, eight, and ten years after entering college are $36,204, $41,459, and $46,090, respectively (in 2014 dollars). The median earning is lower than the average, being $32,922, $37,248, and $40,799, respectively. In the analysis sample, average earnings are $36,626, $41,980, and $46,700, which is about $400-$600 higher than that in the full sample, or by roughly 1 to 1.5%. The numbers suggest that students graduating from institutions that have changed their levels or have reported extremely high or low values of spending per student earn relatively less, though the difference is small.

In Table 2, I present summary statistics for the treatment variable and the set of control variables. First, let us focus on instructional spending. The average instructional spending per student in the full sample is $9,722. The extremely high entries of spending lead to not only a higher level of average spending, even after being weighted by undergraduate enrollments, but also an implausibly large standard deviation of $129,516. In the analysis sample, average instructional spending per student is $8,066, with a much more reasonable standard deviation of $5,935.[15] In Figure 4, I plot the kernel density of the log instructional

does not have an enormous impact on the population of general interest.

[13]Not all institutions appear in all years.

[14]As noted in Section 2.1, I average over an additional year to check for robustness due to the slight misalignment between fiscal year and academic year. Results are almost identical and are not presented.

[15]I compare $8,066 (in 2014 dollars) to similar statistics found from Table 373 in the Digest of Education Statistics 2010, published by the National Center for Education Statistics. In the Digest, instructional spending per student from AY2003-04 to AY2007-08 ranges between $7,792 and $8,221, after adjusting to 2014 dollars. The average $8,066 lies perfectly within the range, while $9,722 is unreasonably too high. I note the difference that my average is calculated over AY1996-97 to AY2007-08, a more extended period

spending in the full sample and the analysis sample. The log values of the two samples are centered at the same point. A small amount of the mass in the tails in the full sample is moved closer to the mean in the analysis sample, as I exclude extreme values. Overall, they align nicely.

The first set of control variables in the analysis sample is measured for the entering cohort. They include the average SAT score (1,076), the average family income ($40,244), the average median household income ($82,677), and the average age at entry (23.9). Another set of the control variables are demographic characteristics averaged over the region where students of the entering cohort come from. They include the percent of White, Black, Asian, and Hispanic from students' zip codes through the Census data, as well as the percent holding Bachelor's degree, Graduate or Professional degree, born in the U.S., and the local unemployment rate and poverty rate. I also control for characteristics of all the enrolled students in the institution, namely, the fraction that is female, is married, is dependent, is veteran, and is a first-generation college student. For all control variables, the analysis sample has almost identical means to the full sample.

From the above comparisons, the analysis sample is representative of the characteristics of postsecondary institutions and the college students in the United States. It covers 5,360 institutions, representing 89% of all undergraduate enrollment during its period. To avoid losing observations due to missing values resulting from item non-response, I recoded the missing values of the control variables to negative ones. I then added indicator variables that equal one if the corresponding data element is missing and zero otherwise.

3 Model and Specification

3.1 Model

I consider the relationship between spending and student earning as a production process, as discussed in Black and Smith (2006). Following their formulation, consider the education production function

than that reported in the Digest.

$$Y = f(x_1, ..., x_k, U, \varepsilon), \tag{1}$$

where Y denotes student log earnings. I denote different educational inputs such as spending, peer quality, student-to-faculty ratio and others with $x_1, ..., x_k$. U represents other observed factors affecting student earnings, and ε is an idiosyncratic error.

I am interested in estimating the parameter

$$\beta = \frac{\partial E(Y|x_1, ..., x_k, U)}{\partial x_1}, \tag{2}$$

where x_1 is the instructional spending per student. This parameter is of particular interest to policymakers in making spending decisions, especially when facing budget shocks.

It is implausible to argue that one can exhaust the list and control for all educational inputs $x_1, ..., x_k$, nor do I make this claim. Black and Smith (2006) lament that most existing empirical work that (implicitly) claims to have estimated the parameter β only controls for a single input x_1. They point out that controlling for only x_1 leads to the estimation of a different parameter

$$\beta' = \frac{\partial E(Y|x_1, U)}{\partial x_1}, \tag{3}$$

that is of no clear economic interpretation and empirical relevance. In addition, omitted variable bias arises when the researcher fails to control for another input x_j that is correlated with x_1. For instance, they showed that five measures or proxies of input (faculty-student ratio, rejection rate, freshman retention rate, mean SAT score, and mean faculty salaries) are positively correlated with the coefficient of correlation ranging from 0.3 to 0.7. Therefore, β' will likely pick up part of the effects of other inputs and hence be greater than β. Indeed, I find similar correlations in my data. Using three variables from my data: average SAT score of the entering class, instructional spending per student, and first-year retention rate, I demonstrate their correlation coefficients in Table 3. In the top panel, they range from 0.4 to 0.7, very close to what was found in Black and Smith (2006).

I make two improvements that significantly ameliorate this issue. First, I control for another input: the average SAT score of the entering class. Consider the first-year retention rate as an omitted variable. In the top panel of Table 3, the correlation coefficient between instructional spending and retention rate is 0.45. I show in the mid-panel that conditional on the average SAT score, the remaining variation in instructional spending and retention rate has a correlation coefficient of only 0.03. Controlling the average SAT score could greatly reduce bias caused by the omission of other inputs that are correlated with the SAT score. Second, due to the panel nature of my data, any inputs that are not time-varying are controlled by including institution fixed effects, even if unobserved or unquantifiable, such as the location of the institution. For time-varying inputs, again, consider the first-year retention rate as the omitted variable. The retention rate and instructional spending are both correlated to the level of prestigious of the institution, which is unobserved and cannot be controlled. By subtracting from the two variables their institution average, in the bottom panel of Table 3, the correlation coefficient between the demeaned values is only 0.05, much smaller than 0.45 in the top panel. Therefore, including institution fixed effects could also greatly reduce the potential bias caused by time-varying inputs that are not controlled in the regression.

Therefore, I consider the parameter I estimate in the reduced form a close approximation to the parameter of interest β.

3.2 Two-way Fixed Effects Regression

I estimate the following reduced form equation using the analysis sample defined in Section 2.

$$\ln(Earning)_{ic} = \alpha + \beta \cdot \ln(InExp)_{ic} + \boldsymbol{X_{ic}} \cdot \boldsymbol{\gamma} + \boldsymbol{\kappa_i} + \boldsymbol{\eta_c} + \boldsymbol{\theta_{s(i)}} \cdot c + u_{ic} \tag{4}$$

where for institution i cohort c, $\ln(Earning)_{ic}$ is the log of one of the earning outcomes and $\ln(InExp)_{ic}$ is the log of the average instructional spending per student defined in Section 2.2. $\boldsymbol{X_{ic}}$ denotes the set of control variables described in Section 2.3. In addition, I also control for the program offering in each institution by including the share of degrees awarded to each Classification of Instructional Programs (CIP) code.

Exploiting the advantages of this panel data set, $\boldsymbol{\kappa_i}$ controls for the institution fixed effects. It is well understood that estimating using cross-sectional data without controlling for institution fixed effects is likely to suffer from omitted variable bias as I discussed in

Section 3.1.

School spending and student earnings are likely to be simultaneously affected by business cycles. I include $\boldsymbol{\eta_c}$ as controls for the cohort fixed effects and $\boldsymbol{\theta_{s(i)}} \cdot c$ as controls for linear state-specific time trends to address the issue. The cohort fixed effects control for cohort-specific events at the national level, such as recessions. But that alone may not be enough, as labor market conditions can vary differently across the nation. I include the state-specific time trends to better address this problem. I do so at the state level for the following two reasons. First, for public institutions, spending changes are primarily affected by state legislation and state appropriations, which happens at the state level. Second, the time trends should ideally capture the movement in economic conditions in different labor markets. Each state can be viewed as a large and relatively concentrated labor market, though not perfectly isolated. It is inappropriate to consider each institution as a separate labor market, and hence I do not replace state-specific time trends with institution-specific time trends. The error term u_{ic} is assumed to be independent of $\ln(InExp)_{ic}$ after conditioning on all the control variables, fixed effects, and the state-specific time trends.

I allow for arbitrary correlation in error terms within an institution by conducting inference using institution-clustered standard errors. If not taken into account, such correlations often vastly underestimate the standard errors of the estimated coefficients, leading to over-rejection of the null hypothesis. As a robust exercise, I also calculate the cluster-robust bootstrap standard errors that maintain the error structure within an institution by resampling at the institution level.

I start by estimating a regression that does not exploit the panel feature of my data. Pooling all the institution-cohort observations, I regress $\ln(Earning)_{ic}$ on $\ln(InExp)_{ic}$ and the controls only to mimic the estimation of early studies. I then add the fixed effects and eventually the state-specific time trends to see how the estimated coefficients evolve. The preferred specification is the complete Equation 4. Lastly, I interact the treatment with the type of institution to analyze effect heterogeneity.

3.3 Parallel Trends Assumption

The two-way fixed effects specification implicitly makes a parallel trends assumption for identification. Conditioning on all the control variables had institutions that experienced changes in instructional spending not undergone such changes; the change in their average earnings would have behaved in the same way as the institutions that did not experience changes in instructional spending, on average.

In standard difference-in-difference estimation, researchers adopting the parallel trends assumption typically plot a graph showing that the control and treatment groups' average outcomes are two parallel lines before the treatment starts (or in the pre-period). As instructional spending is a continuous variable and continuously changing, it is unclear which institutions are the control group and treated group, and what periods are the pre-period and post-period. I attempt to mimic the idea and provide a figure evaluating whether the parallel trends assumption holds, while I bear in mind that this is at best suggestive. To do so, I regress the instructional spending on all the control variables and predict the residuals, which are my identifying variation. I cut my panel by the midpoint (the year 2002) into two artificial "pre-period" and "post-period".[16] I select the institutions that have a small standard deviation (below the median) in the identifying variation in the "pre-period". Among those institutions, I divide them into three groups: whether the instructional spending increased, decreased, or experienced little to no change ("control group") between 2002 and 2003, and plot the average earnings eight years after entry over time for each group.

The top panel of Figure 8 plots the histogram of the standard deviation in the identifying variation in the "pre-period". The distribution is highly skewed towards zero with a median of 0.103, indicated by the dashed line. Institutions below the median are included for analysis in the bottom panel. In the bottom panel, the red line (increased spending) and the green line (decreased spending) are shifted so that their levels match the blue line (no change, or "control group") in 2002. We see that there is a decreasing pre-trend for institutions with increased spending from 2002 to 2003 compared to the "control group".

[16] Realistically, the changes in treatment happen all the time, and it is not clear what pre-period and post-period should be.

But instead of continuing to decrease, the average earnings increased relative to the "control group" when their spending increased in 2003. The other group of institutions where spending decreased from 2002 to 2003 does not exhibit an obvious pre-trend. There is a slight negative relative difference in earnings to the "control group" in 2003 when the spending decreased.

Although the above evidence is imperfect, it is suggestive that the parallel trends assumption can hold in my data.

3.4 Oster's Sensitivity Analysis

Even with both sets of fixed effects and state-specific time-trends, other time-varying variables may exist that correlate with both spending and earnings. I follow the method first proposed in Altonji et al. (2005) and later formalized in Oster (2019) to analyze the sensitivity of the coefficients estimated using the two-way fixed effects regression to potential confounding factors.

The method involves two sensitivity parameters. The first parameter δ measures the ratio of selection on unobserved confounders to the selection on observed variables. The second parameter R_{max} represents the amount of variation in the outcome variable that would have been explained if all confounding variables had been included. If we choose $R_{max} = 1$, we assume that there is no idiosyncratic variation in student earnings that is uncorrelated with instructional spending.[17] I describe the approach in more detail in Appendix A.

With a choice of R_{max}, a researcher can ask two questions under this framework. First, how "strong" do the confounding factors need to be relative to the existing control variables in bringing the estimated coefficients toward zero? A "strong" confounding factor is one that highly correlates with the treatment variable. Second, for a given value of δ, what will the estimate of β be after adjusting for such level of confounding?[18] An intuitive choice

[17] By choosing $R_{max} = 1$, we assume that all the variation in student earnings can be explained by either (1) instructional spending or (2) observed control variables including the fixed effects and state-specific time trends, or (3) unobserved variables that are orthogonal to the observed control variables and are correlated to instructional spending (confounders).

[18] Oster assumes the case where controlling for the unobserved confounders will cause the estimated

considered in Altonji et al. (2005) is $\delta = 1$, which represents the case where the selection on unobserved confounders is equally important as the selection on observed variables.

A choice of R_{max} also has to be made. Oster (2019) studies 76 results in 27 papers published in top journals. She finds that a choice of R_{max} following the rule $R_{max} = \min(1.3\tilde{R}, 1)$ will allow 90 percent of the results using randomized data to survive a $\delta = 1$. The $\tilde{R}$ here is the R-squared in the controlled regression. In my case, it is the within R-squared in Equation 4. Choosing R_{max} using this rule has an intuitive understanding. When choosing control variables, researchers usually choose the ones that are most relevant to explaining the outcomes according to either theory or previous knowledge. Therefore, adding the remaining confounders may only explain a small additional proportion of the variation in the outcome variables. I follow Oster (2019) and set the $R_{max} = \min(1.3\tilde{R}, 1)$ in the analysis.

4 Empirical Results

4.1 Pooled OLS Regression Results

Table 4 presents the pooled OLS regression results. The estimates in column (1) suggest that the elasticities of instructional spending on earnings are 2.7, 3.5, and 4 percent for mean earnings six, eight, and ten years after entering college. Those are comparable in magnitude to other estimates in the literature using cross-sectional data. In column (2), I include only one additional control variable: the average SAT score of the entering class, while not exploiting the panel feature of my data. The estimated coefficients are brought down by a third. It implies that including only one educational input while not controlling for others can lead to biased estimates as many educational inputs are positively correlated. In the third column, I additionally include both the cohort fixed effects and institution fixed effects. The estimates for mean earnings eight and ten years after entry are further brought down by 30-40 percent, becoming 1.7 and 1.9 percent. The estimate for mean earning six years after entry is brought down by two-thirds and becomes insignificant. It is evident

coefficient to move in the same direction as controlling for the observed variables does. In this paper, they move towards zero.

that using cross-sectional data and not controlling for the fixed effects also produces biased estimates. Lastly, in column (4), I use the full specification by including the state-specific time trends, and the estimates are not different from those in column (3).

4.2 Two-way Fixed Effects Regression Results

In Table 5, I present the results for all outcome variables using the specification described in Equation 4. I also summarize the estimated coefficients along with the confidence intervals in Figure 5 to visually present the findings. We observe positive estimated coefficients for all earning outcomes, though the coefficients for the earning outcomes six years after entry are not all statistically significant.

The first observation is that the effects are not obvious shortly after graduation (six years after entry) but become more obvious eight and ten years after entry. The magnitude of the estimated elasticities is small, being 0.5, 1.7, and 1.9 percent for mean earnings six, eight, and ten years after entry. To put the numbers into perspective, the mean value of instructional spending in my analysis sample is $8,066. The average earnings ten years after entry is $46,700. If spending increases by 10%, that is, by $806, it will lead to a 0.19% increase in earnings ten years after entry, which is $89.

The second observation is that the estimated elasticities are higher at lower percentiles in the earning distribution. This pattern holds for earnings six, eight, and ten years after entry. It suggests that higher instructional spending is overall effective in improving student earnings, and more so for students of lower earnings. However, since the estimates are elasticities, the effects in levels do not necessarily obey this pattern.

Different earning outcomes use slightly different sets of institutions due to data availability. To ensure the differences in results are not driven by the sample variation, Table 6 presents the results for the same specification using the same set of institutions where all earning outcome variables are available. The pattern that estimated elasticities are higher at the lower percentiles in the earning distribution remains the same. The point estimates are 1.1, 1.3, and 1.1 percent for earnings six, eight, and ten years after entry. They are not statistically different from previous results using the larger set of institutions. However, they do differ in magnitude and are less dispersed, suggesting that the differences between the estimates for earnings six, eight, and ten years after entry could be caused by differential samples.

4.3 Results by Types of Institution

A natural question to ask is if the above-estimated effects differ by different types of institutions. As the data cover more than 5,300 institutions, I have the statistical power to estimate those effects separately. I interact the instructional spending with an indicator for public institutions. Doing so allows a different slope coefficient to be estimated for public and private institutions.

The results are presented in Table 7. It becomes apparent that the estimated elasticities are different for public and private institutions. The coefficients are 0.6, 2, and 2.3 percent for mean earning six, eight, and ten years after entry for private institutions. For public institutions, they are only 0.3, 0.6, and 0.35 percent and are statistically insignificant. I again visually summarize the coefficients and the confidence intervals for private institutions in Figure 6. The overall pattern looks similar between Figure 5 and Figure 6 that the coefficient is larger at lower earning percentiles.

It warrants further investigation of why instructional spending is effective in private institutions but ineffective in public institutions. I divide the public institutions into more established ones and less established ones. I make the distinction by whether they were observed in 1990. Only a tiny fraction of the institutions that were not observed in 1990 are recently established. As reporting was not mandatory for those not receiving Title IV funds, the majority of them chose not to report to the IPEDS in 1990 and decided to report later on. Table 8 presents the results for public institutions by whether they were observed in 1990. Estimated coefficients for institutions that are more established are 1.1, 1.3, and 0.8 percent. There are closer to the estimated coefficients for private institutions and are statistically significant. The fact that less established institutions face a higher marginal cost in the market of college professors may explain this difference. They usually have to offer higher compensation to attract faculty members of equal caliber than more established institutions.

I also interact the instructional spending with an indicator for whether the institution is a four-year institution. Table 9 presents the results. The estimated elasticities are 1.1, 3.1, and 3.5 percent for earning six, eight, and ten years after entry for four-year institutions, while they are indistinguishable from zero for two-year-or-less institutions.

4.4 Sensitivity for Coefficients in the Main Results

I present the results of the sensitivity analysis discussed in Section 3.4. Table 10 shows the answer to the first question: for how strong a confounding factor has to be so that the estimated coefficients will be brought to zero? The coefficients and standard errors are the same as reported in Table 5. For results that are insignificant, they obviously can only survive a very low level of confounding. For coefficients that are significant, most of them can survive a level of confounding around $\delta = 0.5$. This is lower than the ideal threshold of $\delta = 1$ for the following reason. As I am including a rich set of control variables, the within variations explained by the controlled regression ($\tilde{R}$) are above 80%, leading to $R_{max} = 1$ when using $R_{max} = \min(1.3\tilde{R}, 1)$. As discussed in Oster (2019), $R_{max} = 1$ is a strict criteria where only 9% of the studies using nonrandomized data can survive a $\delta = 1$. To back up the above point, my coefficients for earnings ten years after entry can survive a $\delta = 1$ when the within variation explained by the controlled regression is not as high (around 70%) and $R_{max} = \min(1.3\tilde{R}, 1) < 1$.

The second question is: for a given value of δ, what will the estimate of β be after adjusting for that level of confounding? To answer it, Figure 7 plots the estimated coefficient against values of δ ranging from 0.1 to 1, for all outcome variables. The plot ends when the next value of δ will change the sign of the estimated β. It happens at the value of the corresponding δ shown in Table 10. It is worth noting that, for the elasticity of the mean earnings ten years after entry, controlling for an unobserved confounder that is half as strong as all currently controlled variables will only bring the estimated coefficients down from 1.9 percent to 0.9 percent. From Table 5, the standard error is 0.23 percent, indicating that the coefficient will remain positive and likely remain significant. The same holds for all the elasticities of earning outcomes ten years after entry in the distribution.

5 Instrumental Variable (IV) Regression

In this section, I use the variable ($Z_{i,c}$) constructed by Deming and Walters (2017) as an instrument variable for the instructional spending. Specifically,

$$Z_{i,c} = \left(\frac{Approp_{i,90}}{TotalRevenue_{i,90}}\right) \cdot \left(\frac{StateApprop_{s(i),c}}{FTEStudent_{s(i),c}}\right) \tag{5}$$

where for institution i, $Approp_{i,90}$ is the state appropriation in 1990 and $TotRev_{i,90}$ is the total revenue in 1990. The first factor measures institution i's historical financial dependence on state appropriations. In the second factor, $s(i)$ denotes the state of the institution i. Therefore, the second factor is the state-level average appropriation per FTE enrolled student in year c. As the state-level appropriation only affects public institutions, I run the IV regression using the set of 1,690 public institutions from my analysis sample. Not all 1,690 public institutions were observed in 1990. Nearly one-fourth of them were not.

In the first stage, I estimate the following equation, including the current and one lag of the instrumental variable, along with the entire set of control variables, institution fixed effects, and cohort fixed effects.

$$\ln(InExp_{ic}) = \alpha_1 + \beta_{11}\ln(Z_{i,c}) + \beta_{12}\ln(Z_{i,c-1}) + \boldsymbol{X_{ic}\gamma_1} + \boldsymbol{\kappa_{1i}} + \boldsymbol{\eta_{1c}} + u_{1ic} \tag{6}$$

I do not include state-specific time trends, which deviates from my main specification in the two-way fixed effects estimation. I do so because the instrument is largely affected by state-level time trends. The second factor $\left(\frac{StateApprop_{s(i),c}}{FTEStudent_{s(i),c}}\right)$ only varies at the state level. Including state-specific time trends will absorb almost all variations in the instrument. I include both the contemporaneous and one lag of $Z_{i,c}$ to have maximum predictive power in the first stage. The F-statistic for the joint test of the significance of the contemporaneous and one lag of the instrument is 18.6.

I recalculate the two-year or four-year average instructional spending based on the institution level with the predicted single-year instructional spending. Using that, in the second stage, I estimate the following equation

$$\ln(Earning_{ic}) = \alpha_2 + \beta_2 \ln(In\hat{E}xp_{ic}) + \boldsymbol{X_{ic}\gamma_2} + \boldsymbol{\kappa_{2i}} + \boldsymbol{\eta_{2c}} + u_{2ic} \tag{7}$$

As the two stages are separately estimated, I cluster-bootstrap the entire process to calculate the standard errors.

Table 11 presents the IV regression results. We observe that the estimated coefficients bounce around zero and are all imprecisely estimated, with the standard errors five to ten times as large as those from the two-way fixed effects regressions. This is unsurprising given that we find small to no effect for public institutions in Section 4.3. The IV regression results do not further support, nor contradict the two-way fixed effects results. Unfortunately, the IV regression does not provide additional insights.

6 Discussion

Although this paper intends to inform policymakers in making decisions on adjusting instructional spending when facing budget shocks, in no way do I suggest that a conclusive decision can be made merely based on the results presented in this paper. I investigated earning outcomes, which is only one of the outputs of the education production function. Potentially, the same set of inputs that promote the earnings margin also improve many other margins. For example, higher spending may increase the possibility of going to graduate school, the possibility of working in a job that brings less disutility,[19] and the possibility of meeting a better spouse. Those dimensions are no less important than monetary returns and await further empirical investigations.

To put my estimated coefficients into perspective, I evaluate them at my analysis sample's mean earnings and instructional spending. A standard deviation increase in instructional spending ($5,935) from the mean ($8,066) leads to increases in earnings eight years after entry by $601 overall ($687 for private institutions and $447 for public institutions that are more established). This is comparable to the estimates in Mountjoy and Hickman (2021) where they find a standard deviation increase in a composite index of institution

[19]In Griffith and Rask (2016), they find a marginally significant positive association between higher instructional spending and the probability that one works a job matching the field of study in college.

quality increases student earnings eight to ten years after graduation by $753. Griffith and Rask (2016), however, find a larger elasticity coefficient. According to their estimate, such an increase will lead to an increase in earnings of $1,333 four years after graduation, with a wide confidence interval covering zero.

Though I find instructional spending effective in promoting student earnings, is it still effective considering the cost? To answer this question, I plot the net present return of an investment in increasing instructional spending from the sample mean for four years in a four-year institution. I consider three different discount factors, representing different levels that a person values a dollar one year from now. The first is based on the nominal interest rate in Oct 2021, which is 0.25% so that the discount factor is 0.9975. The second is based on the nominal interest rate in Sep 2019 to be free from the drastic changes in economic conditions due to the pandemic, and the discount factor is 0.9825.[20] The third is based on the real interest rate in Sep 2019, and the discount factor is 1.0006.[21] The first and the third choices of the discount factors reflects a highly forward-looking perspective due to the recent economic conditions in the United States. The third choice corresponds to a view that values a dollar in the future more than a dollar now. The cumulative return assumes the estimated elasticities (in Table 5, Table 7, and Table 8) for earnings eight years after first attendance are persistent throughout students' lifetime.[22]

Figure 9 plots the net present return as the cumulative net present value divided by total investment. The total investment is an increase in instructional spending every year for four years.[23] Using the discount factor 0.9975, a student has to work for 42 years (37 years for private institution and 58 for public institution attendees) until the net present return becomes non-negative. Using the less forward-looking discount factor 0.9825, a private institution attendee has to work for 58 years (or to the age of 80, assuming graduating at 22) until the net present return becomes positive. Using the third discount factor 1.0006, where the student values future more than current, she has to work for 39 years (35 years for

[20]The monthly average of the 1-year treasury bill secondary market rate is 1.75%, retrieved from the Federal Reserve Bank.

[21]Annual inflation rate is 1.81% in 2019, retrieved from the Bureau of Labor Statistics.

[22]The combined evidence in the literature supports the assumption that the elasticities can be persistent.

[23]I consider four-year institutions only because in Section 4.3, only four-year institutions were found to have positive effects.

private and 53 years for public institution attendees) until the investment in instructional spending start to payoff.

Alternatively, I evaluate for what value the discount factor has to take so that the net present return is non-negative after working for a fixed number of years. Figure 10 plots such relationships for overall, private, and public institutions assuming a student works for 35 to 55 years. First, public institutions require an extremely forward-looking perspective, given its low estimated return. In other words, unless one views future dollars more than current dollars, investment in instructional spending cannot payoff. For example, if one works for 35 years, a discount factor of 1.02 is required, which is unrealistic. For private institutions, up to a small degree of discounting, investment in instructional spending can still payoff. If one works for 45 years, a discount factor of 0.9889 will break even.

The results above may justify that institutions have been shifting away from allocating additional funds to instructional spending. However, the budget decision process surely do not solely rely on the perceived effectiveness of instructional spending, but is oriented by the values and goals of different institution. Future research is warranted to evaluate how institutions make their budget decisions.

References

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of political economy 113*(1), 151–184.

Behrman, J. R., M. R. Rosenzweig, and P. Taubman (1996). College choice and wages: Estimates using data on female twins. *The Review of Economics and Statistics*, 672–685.

Black, D. A. and J. A. Smith (2006). Estimating the returns to college quality with multiple proxies for quality. *Journal of labor Economics 24*(3), 701–728.

Dale, S. and A. B. Krueger (2011). Estimating the return to college selectivity over the career using administrative earnings data. Technical report, National Bureau of Economic Research.

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*(4), 1491–1527.

Deming, D. J. and C. R. Walters (2017). The impact of price caps and spending cuts on us postsecondary attainment. Technical report, National Bureau of Economic Research.

Foster, E. and J. Rodgers (1980). Quality of education and student earnings. *Higher education 9*(1), 21–37.

Griffith, A. L. and K. N. Rask (2016). The effect of institutional expenditures on employment outcomes and earnings. *Economic Inquiry 54*(4), 1931–1945.

Grogger, J. (1996). School expenditures and post-schooling earnings: evidence from high school and beyond. *The Review of Economics and Statistics*, 628–637.

Hanushek, E. A. (1981). Throwing money at schools. *Journal of policy analysis and management 1*(1), 19–41.

Hoekstra, M. (2009). The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics 91*(4), 717–724.

Hyman, J. (2017). Does money matter in the long run? effects of school spending on educational attainment. *American Economic Journal: Economic Policy 9*(4), 256–80.

Jackson, C. K., R. C. Johnson, and C. Persico (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics 131*(1), 157–218.

Lafortune, J., J. Rothstein, and D. W. Schanzenbach (2018). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics 10*(2), 1–26.

Mountjoy, J. and B. R. Hickman (2021). The returns to college(s): Relative value-added and match effects in higher education. Technical report, National Bureau of Economic Research.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics 37*(2), 187–204.

Papke, L. E. (2005). The effects of spending on test pass rates: evidence from michigan. *Journal of Public Economics 89*(5-6), 821–839.

Solomon, L. C. (1975). The definition of college quality and its impact on earnings. In *Explorations in Economic Research, Volume 2, number 4*, pp. 537–587. NBER.

Wachtel, P. (1976). The effect on earnings of school and college investment expenditures. *The Review of Economics and Statistics*, 326–331.

Webber, D. A. (2012). Expenditures and postsecondary graduation: An investigation using individual-level data from the state of ohio. *Economics of Education Review 31*(5), 615–618.

Webber, D. A. and R. G. Ehrenberg (2010). Do expenditures other than instructional expenditures affect graduation and persistence rates in american higher education? *Economics of Education Review 29*(6), 947–958.

Table 1: Summary Statistics - Outcome Variables

	(1) Full Sample			(2) Analysis Sample		
	Mean	SD	N	Mean	SD	N
Earning, six years after entry						
Mean	36204	9393	40122	36626	9086	32741
25th percentile	19494	6642	35834	19880	6555	29545
Median	32922	8707	40122	33373	8549	32741
75th percentile	47769	11254	35834	48246	10885	29545
Earning, eight years after entry						
Mean	41459	11562	33803	41980	11175	27535
25th percentile	22385	7788	30240	22859	7701	24866
Median	37248	9997	33803	37782	9830	27535
75th percentile	54199	13559	30240	54771	13094	24866
Earning, ten years after entry						
Mean	46090	13691	27872	46700	13251	22669
25th percentile	24517	8675	24957	25046	8580	20527
Median	40799	11232	27872	41401	11062	22669
75th percentile	60032	16027	24957	60706	15490	20527

Notes: Data from College Scorecard project. The full sample consists of observations where instructional spending and at least one of the earning outcome variables is nonmissing. Analysis sample removes from the full sample (1) institutions that have changed their levels (two-year vs. four-year), and (2) institutions that had higher than $100,000 or lower than $50 instructional spending per student. Mean and standard deviation are weighted by the number of undergraduate enrollments. All monetary terms are in 2014 dollars.

Table 2: Summary Statistics - Treatment and Control Variables

	(1)			(2)		
	Full Sample			Analysis Sample		
	Mean	SD	N	Mean	SD	N
Instructional spending per student	9722	129516	50263	8066	5935	40962
Average SAT	1077	119	11311	1076	115	10668
Family income	38938	19908	50231	40244	19957	40939
Median household income	82071	17743	24716	82667	17681	20102
Age at entry	24.0	2.9	50231	23.9	2.9	40939
White	74.4	15.1	24716	74.6	15.1	20102
Black	12.8	11.8	24716	12.7	11.8	20102
Asian	3.9	4.8	24716	3.9	4.8	20102
Hispanic	13.9	17.4	24716	13.4	16.6	20102
Bachelor	15.6	4.0	24613	15.7	4.0	20028
Graduate or professional	8.7	2.9	24613	8.8	2.9	20028
Born in U.S.	87.4	10.5	24613	87.7	10.2	20028
Unemployment rate	3.77	1.08	24716	3.74	1.05	20102
Poverty rate	10.02	6.00	24716	9.85	5.77	20102
Female	0.60	0.11	45688	0.60	0.10	37290
Married	0.15	0.10	47865	0.15	0.10	38888
Dependent	0.60	0.22	49347	0.62	0.22	40247
Veteran	0.03	0.03	33178	0.03	0.03	27221
First generation	0.43	0.12	47374	0.42	0.12	38703

Notes: Data from College Scorecard project. See notes in Table 1 for sample restrictions. Mean and standard deviation are weighted by the number of undergraduate enrollments. White refers to the percent of the population from students' zip codes who are White, via Census data. The same holds for Black, Asian, Hispanic, Bachelor, Graduate or professional, Born in the U.S., Unemployment rate, and Poverty rate. All monetary terms are in 2014 dollars.

Table 3: Correlation of Different Educational Inputs

	SAT	InExp	Retention
SAT	1		
InExp	0.611	1	
Retention	0.707	0.448	1

	Res_InExp	Res_Retention
Res_InExp	1	
Res_Retention	0.0287	1

	Dm_InExp	Dm_SAT	Dm_Retention
Dm_InExp	1		
Dm_SAT	0.185	1	
Dm_Retention	0.0453	0.112	1

Notes: Data from College Scorecard project. The top panel shows correlation between the raw values of the average SAT score, instructional spending per FTE student, and first-year retention rate. In the mid panel, I correlate residuals of instructional spending and retention rate from two separate regressions on the SAT score, respectively. In the bottom panel, I correlate the demeaned value of the three variables after subtracting from them the institution average.

Table 4: Effect of Instructional Spending on Earnings, Pooled OLS Regression

	(1)	(2)	(3)	(4)
Log mean earning six years after entry				
ln(InExp/FTE)	0.0270	0.0186	0.00545	0.00535
	(0.0053)	(0.0035)	(0.0030)	(0.0027)
Number of Institutions	4721	4721	4721	4721
Log mean earning eight years after entry				
ln(InExp/FTE)	0.0351	0.0271	0.0173	0.0175
	(0.0049)	(0.0038)	(0.0030)	(0.0028)
Number of Institutions	4323	4323	4323	4323
Log mean earning ten years after entry				
ln(InExp/FTE)	0.0403	0.0276	0.0192	0.0194
	(0.0052)	(0.0037)	(0.0023)	(0.0023)
Number of Institutions	3998	3998	3998	3998
Other Controls	Yes	Yes	Yes	Yes
Average SAT	No	Yes	Yes	Yes
Fixed Effects	No	No	Yes	Yes
State-Specific Time Trends	No	No	No	Yes

Notes: Data from College Scorecard project for cohorts between 1996 and 2007. See notes in Table 1 for sample restrictions. Control variables include average SAT score, average family income, median household income, and average age at entry for the entering cohort, racial composition, educational level, unemployment rate and poverty rate of the place of origin of the entering students, the fraction of the enrolled students that is female, is married, is dependent, is veteran, and is first generation college student, and composition of degrees conferred by Classification of Instructional Programs (CIP) codes. Regressions are weighted by the average number of cohort size in each institution. Standard errors in parentheses are calculated using cluster-robust standard errors at the institution level.

Table 5: Effect of Instructional Spending on Earnings

	(1) Mean	(2) Pct25	(3) Pct50	(4) Pct75
Log earnings six years after entry				
ln(InExp/FTE)	0.0053	0.014	0.0057	0.0030
	(0.0027)	(0.0044)	(0.0029)	(0.0026)
Number of Institutions	4721	4721	4721	4721
Log earnings eight years after entry				
ln(InExp/FTE)	0.017	0.025	0.019	0.015
	(0.0028)	(0.0039)	(0.0029)	(0.0027)
Number of Institutions	4323	4323	4323	4323
Log earnings ten years after entry				
ln(InExp/FTE)	0.019	0.031	0.023	0.017
	(0.0023)	(0.0034)	(0.0023)	(0.0021)
Number of Institutions	3998	3998	3998	3998

Notes: Data from College Scorecard project for cohorts between 1996 and 2007. See notes in Table 1 for sample restrictions. All regressions include cohort fixed effects, institution fixed effects, state-specific time trends and a set of control variables. Outcome variables are mean, 25th percentile, median, and 75th percentile of earningss six, eight, and ten years after entry. Control variables include average SAT score, average family income, median household income, and average age at entry for the entering cohort, racial composition, educational level, unemployment rate and poverty rate of the place of origin of the entering students, the fraction of the enrolled students that is female, is married, is dependent, is veteran, and is first generation college student, and composition of degrees conferred by Classification of Instructional Programs (CIP) codes. Regressions are weighted by the average number of cohort size in each institution. Standard errors in parentheses are calculated using cluster-robust standard errors at the institution level.

Table 6: Effect of Instructional Spending on Earnings, Same Institutions

	(1) Mean	(2) Pct25	(3) Pct50	(4) Pct75
Log earnings six years after entry				
ln(InExp/FTE)	0.011	0.017	0.011	0.0087
	(0.0024)	(0.0036)	(0.0024)	(0.0023)
Log earnings eight years after entry				
ln(InExp/FTE)	0.013	0.022	0.014	0.010
	(0.0023)	(0.0030)	(0.0023)	(0.0023)
Log earnings ten years after entry				
ln(InExp/FTE)	0.011	0.019	0.014	0.0092
	(0.0020)	(0.0029)	(0.0020)	(0.0019)
Number of Institutions	3826	3826	3826	3826

Notes: Data from College Scorecard project for cohorts between 1996 and 2007. See notes in Table 1 for sample restrictions. All regressions include cohort fixed effects, institution fixed effects, state-specific time trends and a set of control variables. Outcome variables are mean, 25th percentile, median, and 75th percentile of earningss six, eight, and ten years after entry. Control variables include average SAT score, average family income, median household income, and average age at entry for the entering cohort, racial composition, educational level, unemployment rate and poverty rate of the place of origin of the entering students, the fraction of the enrolled students that is female, is married, is dependent, is veteran, and is first generation college student, and composition of degrees conferred by Classification of Instructional Programs (CIP) codes. Regressions are weighted by the average number of cohort size in each institution. Standard errors in parentheses are calculated using cluster-robust standard errors at the institution level.

Table 7: Effect of Instructional Spending on Earnings, by Public and Private Institutions

Treatment: Ln(InExp/FTE)	(1)	(2)	(3)	(4)
	Mean	Pct25	Pct50	Pct75
Log earnings six years after entry				
Private	0.0059	0.016	0.0056	0.0030
	(0.0032)	(0.0051)	(0.0033)	(0.0030)
Public	0.0030	0.0068	0.0065	0.0034
	(0.0044)	(0.0066)	(0.0048)	(0.0045)
Number of Institutions	4721	4721	4721	4721
Log earnings eight years after entry				
Private	0.020	0.029	0.021	0.017
	(0.0032)	(0.0045)	(0.0033)	(0.0030)
Public	0.0060	0.0068	0.0079	0.0047
	(0.0053)	(0.0065)	(0.0055)	(0.0051)
Number of Institutions	4323	4323	4323	4323
Log earnings ten years after entry				
Private	0.023	0.037	0.027	0.020
	(0.0025)	(0.0039)	(0.0026)	(0.0024)
Public	0.0035	0.0048	0.0036	0.0012
	(0.0040)	(0.0050)	(0.0036)	(0.0037)
Number of Institutions	3998	3998	3998	3998

Notes: Data from College Scorecard project for cohorts between 1996 and 2007. See notes in Table 1 for sample restrictions. All regressions include cohort fixed effects, institution fixed effects, state-specific time trends and a set of control variables. See notes in Table 4 for descriptions of the control variables. Treatment variable is instructional spending per FTE student. Instructional spending is interacted with the type of the institution. Outcome variables are mean, 25th percentile, median, and 75th percentile of earnings six, eight, and ten years after entry. Regressions are weighted by the average number of cohort size in each institution. Standard errors in parentheses are calculated using cluster-robust standard errors at the institution level.

Table 8: Effect of Instructional Spending on Earnings, Public Institutions

Treatment: Ln(InExp/FTE)	(1)	(2)	(3)	(4)
	Mean	Pct25	Pct50	Pct75
Log earning six years after entry (No. of Institutions: 1686)				
Not Observed in 1990	-0.0063	-0.0016	-0.0059	-0.0087
	(0.0050)	(0.0064)	(0.0054)	(0.0051)
Observed in 1990	0.011	0.012	0.014	0.012
	(0.0037)	(0.0053)	(0.0040)	(0.0037)
Log earning eight years after entry (No. of Institutions: 1648)				
Not Observed in 1990	-0.0012	0.0037	0.00077	-0.0023
	(0.0055)	(0.0075)	(0.0059)	(0.0045)
Observed in 1990	0.013	0.012	0.014	0.012
	(0.0050)	(0.0066)	(0.0055)	(0.0051)
Log earning ten years after entry (No. of Institutions: 1609)				
Not Observed in 1990	-0.0021	0.011	0.0042	-0.0043
	(0.0064)	(0.0084)	(0.0061)	(0.0058)
Observed in 1990	0.0079	0.0094	0.0058	0.0048
	(0.0034)	(0.0044)	(0.0031)	(0.0032)

Notes: Data from College Scorecard project for cohorts between 1996 and 2007. See notes in Table 1 for sample restrictions. All regressions include cohort fixed effects, institution fixed effects, state-specific time trends and a set of control variables. See notes in Table 4 for descriptions of the control variables. Treatment variable is instructional spending per FTE student. Instructional spending is interacted with whether the institution was observed in 1990. Outcome variables are mean, 25th percentile, median, and 75th percentile of earnings six, eight, and ten years after entry. Regressions are weighted by the average number of cohort size in each institution. Standard errors in parentheses are calculated using cluster-robust standard errors at the institution level.

Table 9: Effect of Instructional Spending on Earnings, by Institution Levels

Treatment: Ln(InExp/FTE)	(1) Mean	(2) Pct25	(3) Pct50	(4) Pct75
Log earning six years after entry (No. of Institutions: 4721)				
Two-year or less	-0.0029	-0.0045	-0.0024	-0.00081
	(0.0027)	(0.0044)	(0.0029)	(0.0025)
Four-year	0.011	0.028	0.012	0.0059
	(0.0041)	(0.0068)	(0.0044)	(0.0040)
Log earning eight years after entry (No. of Institutions: 4323)				
Two-year or less	0.0024	0.0024	0.0034	0.0036
	(0.0027)	(0.0038)	(0.0028)	(0.0026)
Four-year	0.031	0.047	0.034	0.025
	(0.0046)	(0.0067)	(0.0048)	(0.0044)
Log earning ten years after entry (No. of Institutions: 3998)				
Two-year or less	0.0052	0.0048	0.0065	0.0054
	(0.0022)	(0.0034)	(0.0023)	(0.0021)
Four-year	0.035	0.059	0.040	0.029
	(0.0036)	(0.0049)	(0.0037)	(0.0034)

Notes: Data from College Scorecard project for cohorts between 1996 and 2007. See notes in Table 1 for sample restrictions. All regressions include cohort fixed effects, institution fixed effects, state-specific time trends and a set of control variables. See notes in Table 4 for descriptions of the control variables. Treatment variable is instructional spending per FTE student. Instructional spending is interacted with whether the institution is a four-year institution. Outcome variables are mean, 25th percentile, median, and 75th percentile of earnings six, eight, and ten years after entry. Regressions are weighted by the average number of cohort size in each institution. Standard errors in parentheses are calculated using cluster-robust standard errors at the institution level.

Table 10: Sensitivity: What value of δ will drive β to 0?

	(1) Mean	(2) Pct25	(3) Pct50	(4) Pct75
Log earnings six years after entry				
ln(InExp/FTE)	0.00535	0.0141	0.00573	0.00303
	(0.0027)	(0.0044)	(0.0029)	(0.0026)
$\tilde{R}$	0.900	0.903	0.917	0.897
$R_{max} = \min(1.3 \cdot \tilde{R}, 1)$	1	1	1	1
δ	0.302	0.508	0.343	0.183
Log earnings eight years after entry				
ln(InExp/FTE)	0.0175	0.0253	0.0191	0.0149
	(0.0028)	(0.0039)	(0.0029)	(0.0027)
$\tilde{R}$	0.806	0.808	0.838	0.801
$R_{max} = \min(1.3 \cdot \tilde{R}, 1)$	1	1	1	1
δ	0.486	0.483	0.561	0.431
Log earnings ten years after entry				
ln(InExp/FTE)	0.0194	0.0310	0.0225	0.0168
	(0.0023)	(0.0034)	(0.0023)	(0.0021)
$\tilde{R}$	0.686	0.662	0.753	0.687
$R_{max} = \min(1.3 \cdot \tilde{R}, 1)$	0.892	0.860	0.980	0.893
δ	0.974	1.139	0.893	0.822

Notes: This table applies the sensitivity analysis described in Oster (2019) to the main estimation equation. The coefficients and the standard errors are the same as in Table 5. $\tilde{R}$ is the R-squared for the controlled regression. R_{max} is conceptually the maximum variation in the outcome variable that could be explained by including all possible confounding variables. Using $R_{max} = \min(1.3 \cdot \tilde{R}, 1)$ was suggested by Oster (2019) based on a review of existing published articles. δ is the degree of selection on potential confounding variables relative to the degree of selection on observed variables that are already included in the equation. Larger δ indicates more robust estimated coefficient.

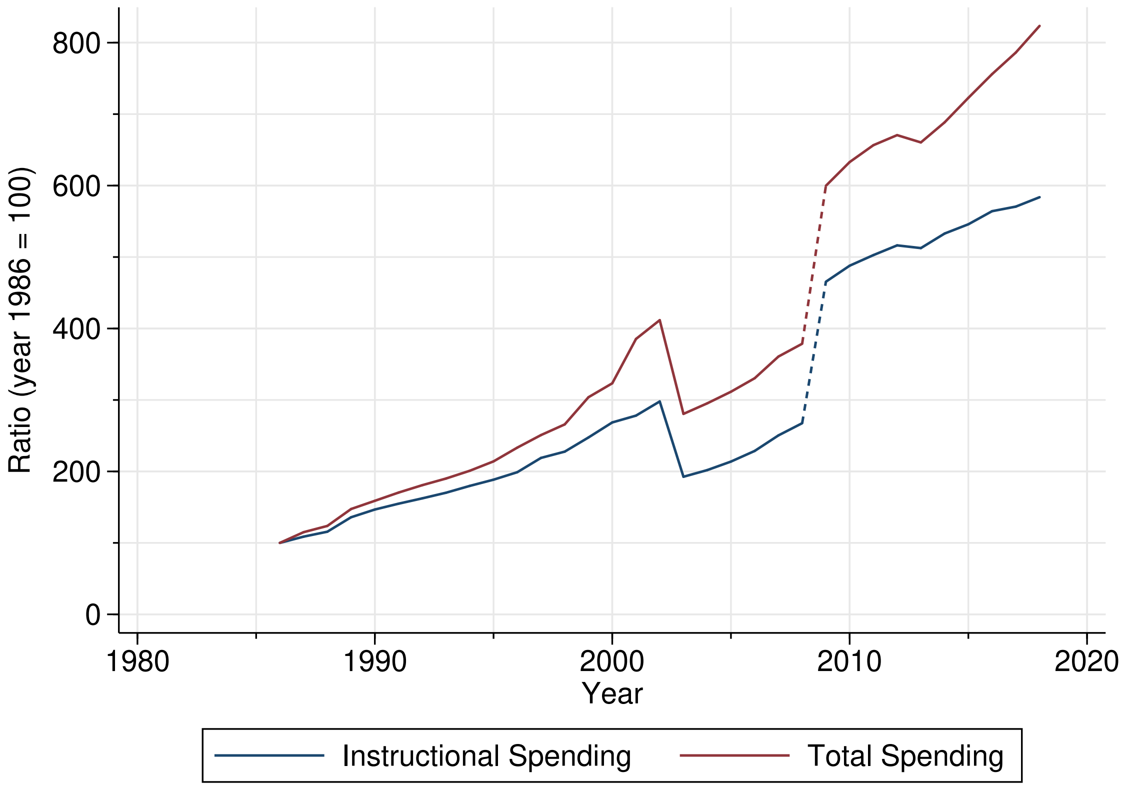
.

Table 11: IV Regression: Effect of Instructional Spending on Earnings, Public Institutions

	(1) Mean	(2) Pct25	(3) Pct50	(4) Pct75
Log earnings six years after entry				
Ln(InExp/FTE)	0.016	-0.041	0.006	0.007
	[.026]	[.044]	[.029]	[.026]
Number of Institutions	1266	1266	1266	1266
Log earnings eight years after entry				
Ln(InExp/FTE)	0.007	-0.066	-0.021	0.008
	[.02]	[.038]	[.025]	[.019]
Number of Institutions	1260	1260	1260	1260
Log earnings ten years after entry				
Ln(InExp/FTE)	-0.009	-0.047	-0.010	0.002
	[.021]	[.03]	[.021]	[.024]
Number of Institutions	1252	1252	1252	1252

Notes: This table presents the IV regression results. The instrument is the log of the institution's financial dependence on state appropriations in 1990 multiplied by average state appropriations per student, and its one-lag value. See Section 5 for more detailed discussion. Data from College Scorecard project and State Higher Education Finance project for cohorts between 1996 and 2007. Analysis sample consists of institution that has remained consistent in its level and has reported instructional spending per student between \$50 to \$100,000. All regressions include a set of control variables, institution fixed effects and cohort fixed effects. See notes in Table 4 for descriptions of the control variables. The first-stage F-statistic in 18.6. Standard errors in brackets are calculated based on 100 replications of cluster-bootstrap of the entire process at the institution level.

Figure 1: Growth of Instructional Spending and Total Spending

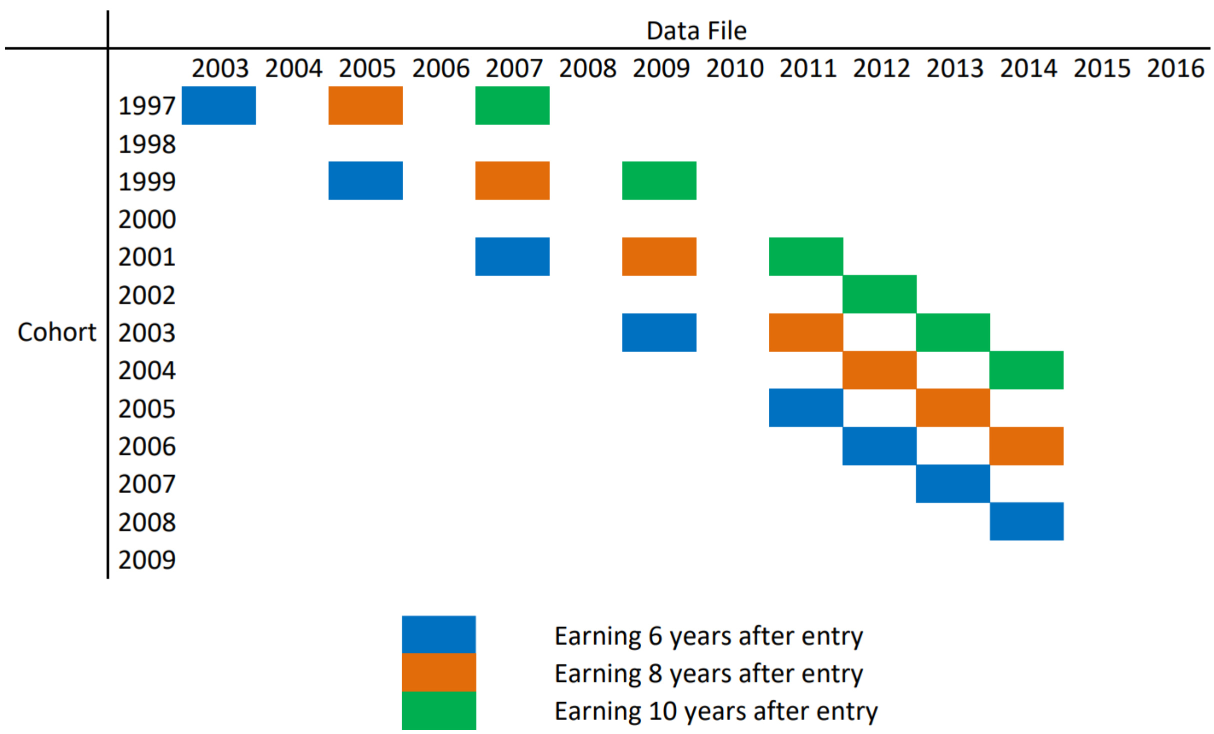


Notes: Author's calculation based on data from the Integrated Postsecondary Education Data System (IPEDS). The dashed line represents an institutional change in data collection process. Comparisons should not be made before and after the change.

Figure 2: Change in Instructional Spending and Change in Wage Earnings of Young Workers with College Degree

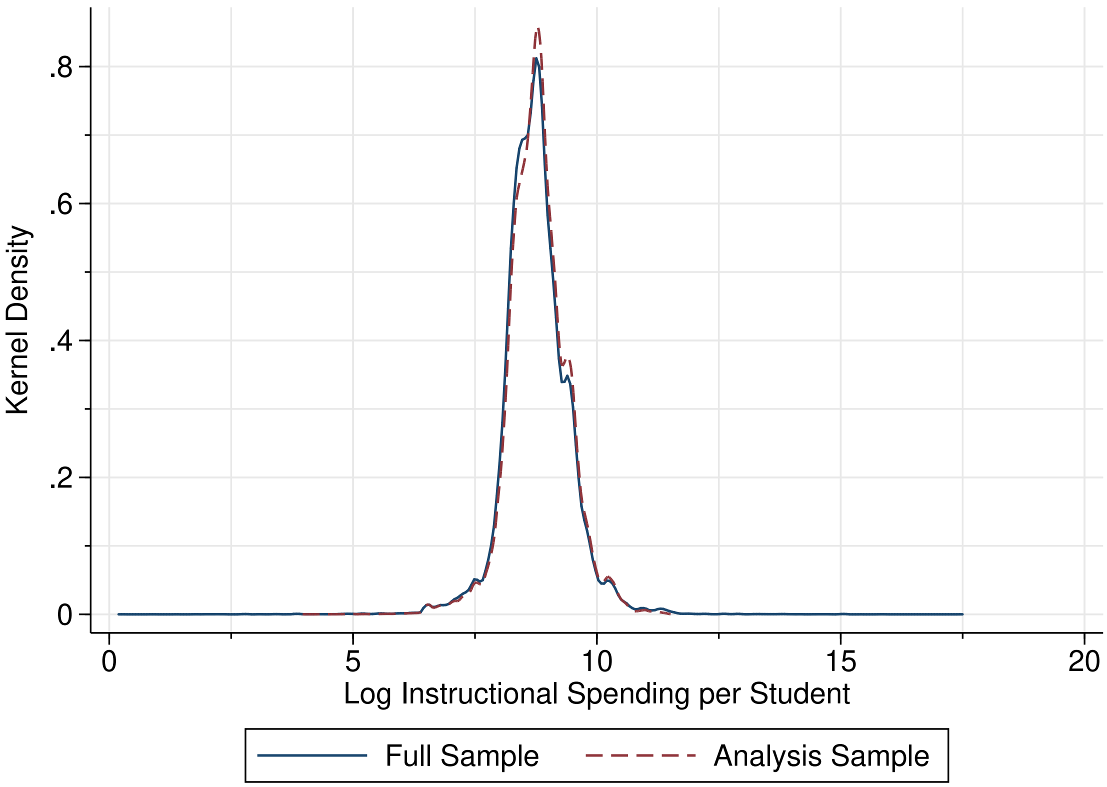
Notes: Each point represent a state. States are grouped into four regions and are indicated by different colors. Horizontal axis shows the percentage change in spending per FTE student between 2009 and 2016. Vertical axis shows the percentage change in wage earnings for young workers (age 22-30) with at least associate degree between 2012 and 2019. The dashed line is a linear fit weighted by the population of young workers with at least associate degree in 2019. Calculations are based on data from Integrated Postsecondary Education Data System (IPEDS) and American Community Survey (ACS). All monetary terms are deflated using Consumer Price Index from Bureau of Labor Statistics.

Figure 3: Data Structure and Availability for Outcome Variables



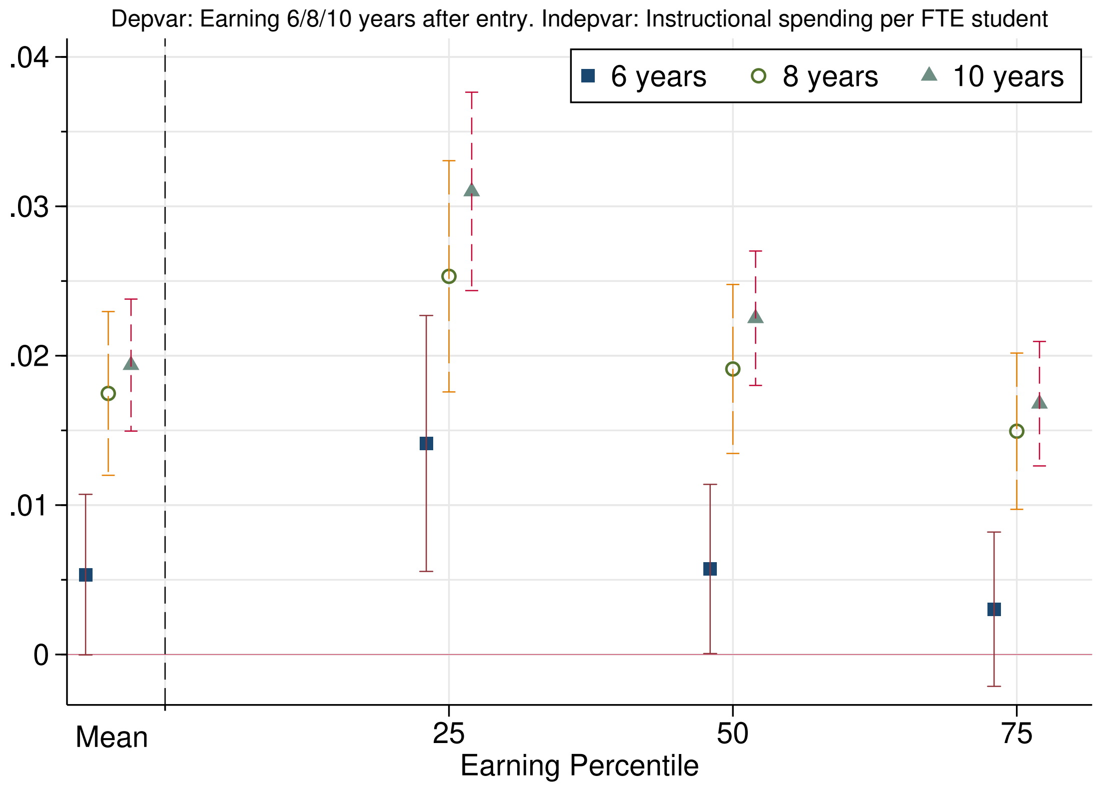
Notes: Colored area indicates that outcome variable is available in that year. Cohorts are pooled at two-years level. Cohort 1997 includes all students entering the institution in Fall 1996 and Fall 1997.

Figure 4: Distribution of the Instructional Spending



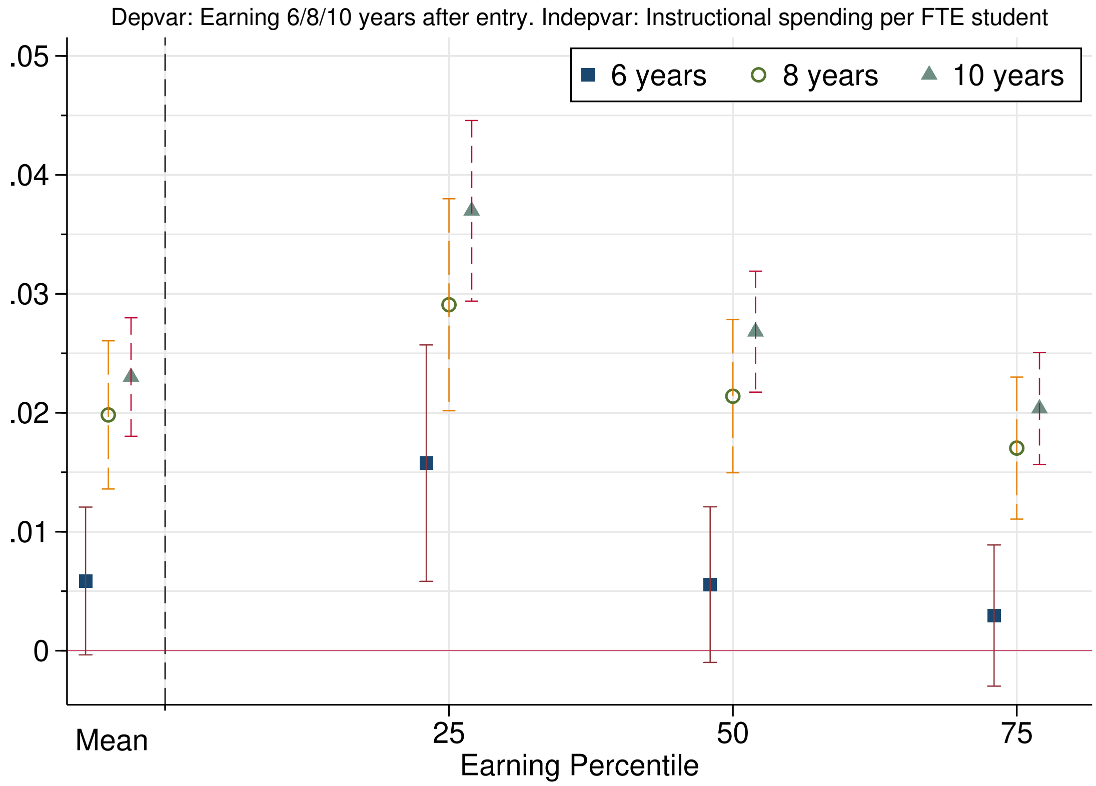
Notes: The full sample consists of 51,779 institution-cohort observations. The analysis sample consists of 42,244 observations, which are institutions that have remained consistent in their levels, and have not reported to have instructional spending per student falling outside the range of $50 to $100,000.

Figure 5: Effect of Instructional Spending per FTE Student on Earnings (All Institutions)



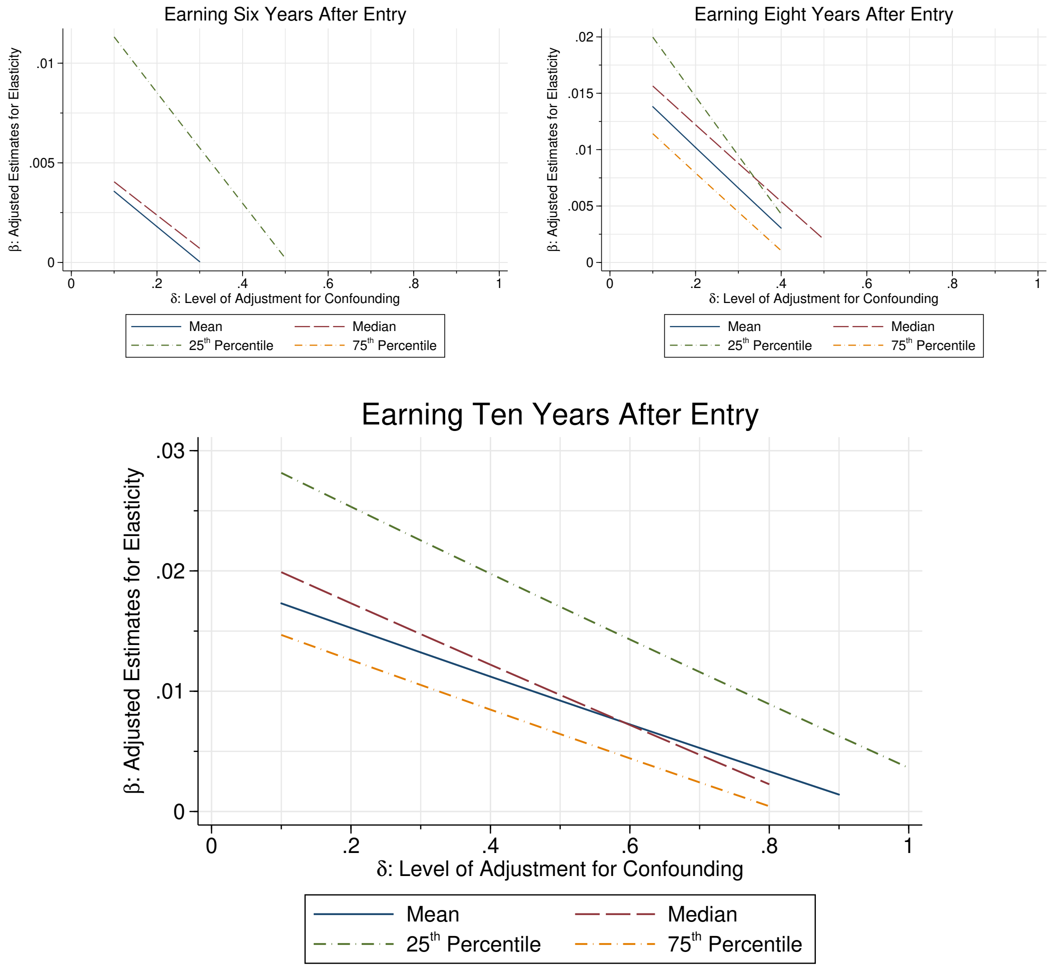
Notes: This figure visually summarizes estimation results of the effect of instructional spending per student on average student earnings. 95% confidence intervals are attached to the point estimates. See Table 5 for the exact values. Data from College Scorecard project for cohorts between 1996 and 2007. See notes in Table 1 for sample restrictions. All regressions include cohort fixed effects, institution fixed effects, state-specific time trends and a set of control variables. Instructional spending is interacted with the type of the institution. Control variables include average SAT score, average family income, median household income, and average age at entry for the entering cohort, racial composition, educational level, unemployment rate and poverty rate of the place of origin of the entering students, the fraction of the enrolled students that is female, is married, is dependent, is veteran, and is first generation college student, and composition of program offering by Classification of Instructional Programs (CIP) codes. Regressions are weighted by the average number of cohort size in each institution. Standard errors are clustered at the institution level.

Figure 6: Effect of Instructional Spending per FTE Student on Earnings (Private Institutions)



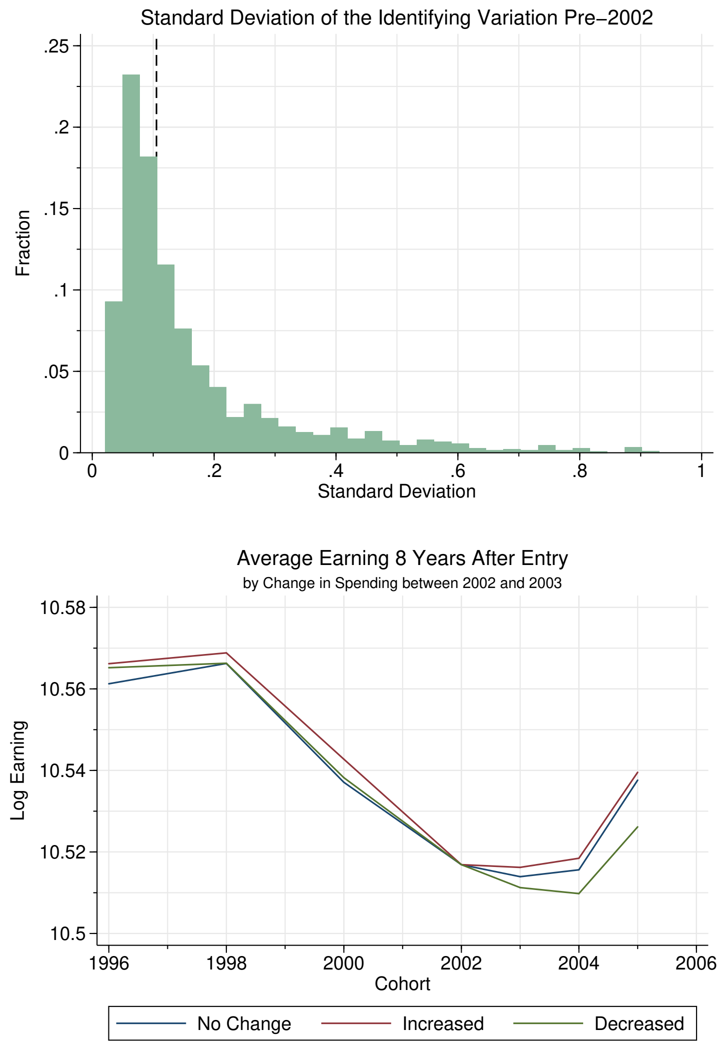
Notes: This figure visually summarizes estimation results of the effect of instructional spending per student on average student earnings for private institutions. 95% confidence intervals are attached to the point estimates. See Table 7 for the exact values. Data from College Scorecard project for cohorts between 1996 and 2007. See notes in Table 1 for sample restrictions. All regressions include cohort fixed effects, institution fixed effects, state-specific time trends and a set of control variables. Instructional spending is interacted with the type of the institution. Control variables include average SAT score, average family income, median household income, and average age at entry for the entering cohort, racial composition, educational level, unemployment rate and poverty rate of the place of origin of the entering students, the fraction of the enrolled students that is female, is married, is dependent, is veteran, and is first generation college student, and composition of program offering by Classification of Instructional Programs (CIP) codes. Regressions are weighted by the average number of cohort size in each institution. Standard errors are clustered at the institution level.

Figure 7: Confounding Adjusted Estimated Elasticities (All Institutions)



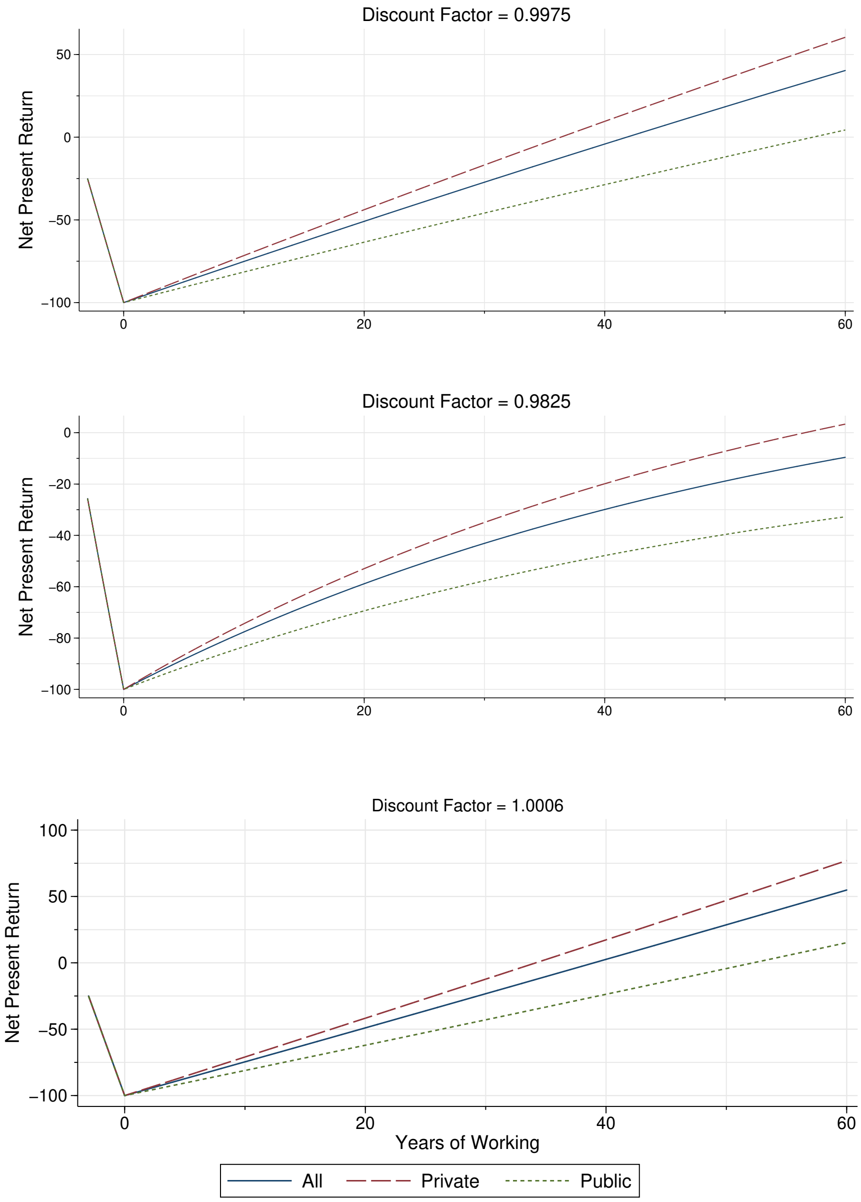
Notes: This figure plots the estimated coefficients in Table 5 against levels of adjustment for confounding factors with the method discussed in Oster (2019). The parameter $R_{\max}$ is chosen as $\min(1.3\tilde{R}, 1)$ where $\tilde{R}$ is the R-squared in the controlled regression, as suggested in the original paper. Data from College Scorecard project for cohorts between 1996 and 2007. See notes in Table 1 for sample restrictions. All regressions include cohort fixed effects, institution fixed effects, state-specific time trends and a set of control variables. See notes in Table 4 for descriptions of the control variables. Regressions are weighted by the average number of cohort size in each institution. Remaining positive for a higher level of confounding adjustment indicates a more robust estimated coefficient.

Figure 8: Parallel Trends Assumption



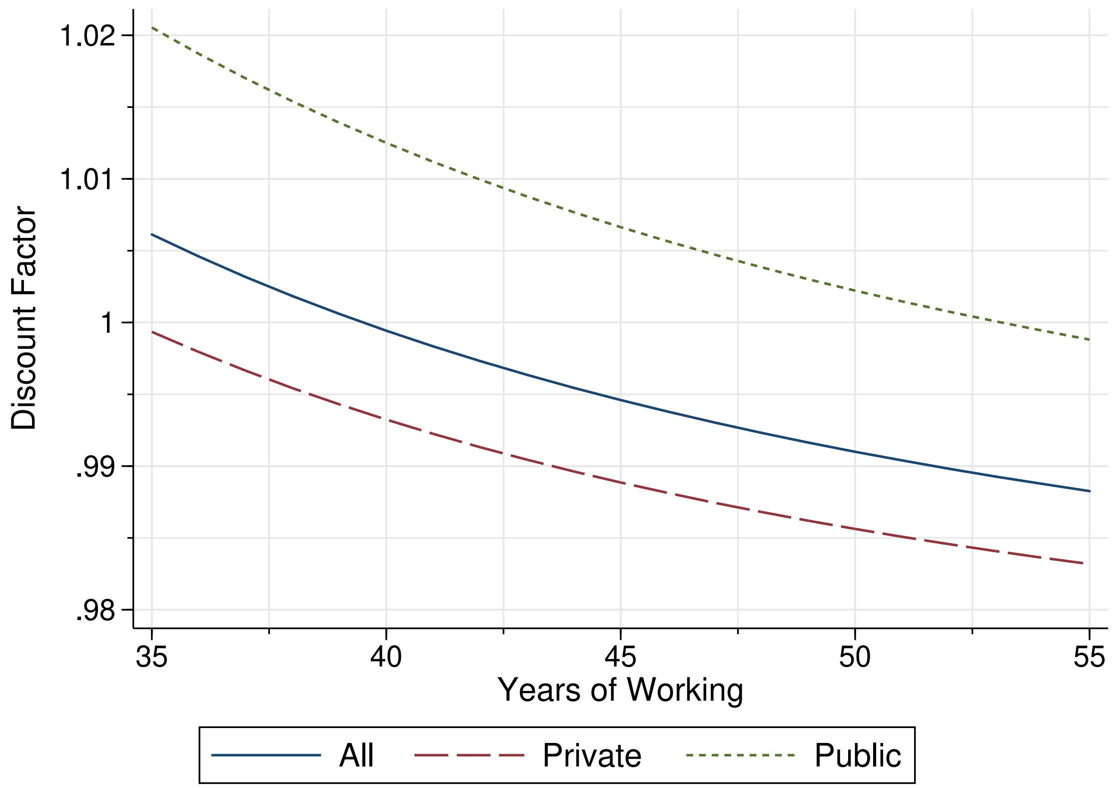
Notes: The figure in the top panel plots the histogram of the standard deviations of the identifying variation in the artificial "pre-period" (before 2002). The identifying variation is the predicted residuals from a regression of instructional spending on all the control variables, cohort fixed effects, institution fixed effects, and state-specific time trends. See notes in Table 1 and Table 4 for detailed descriptions on sample restrictions and the set of control variables. Dashed line indicates the median. Institutions with standard deviations below the median are considered as having little variation in the "pre-period", and are further divided into three groups by whether they experienced increase, decrease, or no change (or little change) in instructional spending between 2002 and 2003. The figure in the bottom panel plots the average earnings eight years after first attendance for the above three groups over time.

Figure 9: Net Present Return of Instructional Spending



Notes: This figure plots the cumulative return of instructional spending as a percentage of total investment in instructional spending at a four-year college. The total investment is an increase in instructional spending each year for four years. The cumulative return assumes the estimated returns (in Table 5, Table 7, and Table 8) for earnings eight years after first attendance are persistent throughout students' lifetime and three values of discount factors: 0.9975, 0.9825, and 1.0006.

Figure 10: Discount Factor for the Investment in Instructional Spending to Break Even



Notes: This figure plots the value of the discount factor so that the net present return of the investment in instructional spending after working for a fixed number of years will break even. The total investment is an increase in instructional spending each year for four years. The cumulative return assumes the estimated returns (in Table 5, Table 7, and Table 8) for earnings eight years after first attendance are persistent throughout students' lifetime.

Appendix

A Oster's Sensitivity Analysis

In this section, I describe the Oster's approach in more detail. Following the formulation in Oster (2019), suppose that the true relationship can be written as

$$\ln(Earning_{ic}) = \alpha + \beta \cdot \ln(InExp_{ic}) + \boldsymbol{W}_{1ic} + W_{2ic} + v_{ic} \tag{8}$$

The term W_{1ic} here corresponds to the collection of observed variables that are controlled in the regression. In my case, $\boldsymbol{W}_{1ic} = \boldsymbol{X_{ic}} \cdot \boldsymbol{\gamma} + \boldsymbol{\kappa_i} + \boldsymbol{\eta_c} + \boldsymbol{\theta_{s(i)}} \cdot c$. The term W_{2ic} represents a combination of all possible unobserved confounders. It also takes care of measurement errors, but not model misspecification. It is assumed that after controlling for W_2, we have $\mathbb{E}(v|X, W_1, W_2) = 0$.

Oster (2019) formalizes the sensitivity analysis first proposed in Altonji et al. (2005). In the framework she described, there are two parameters to be considered. The first parameter $\delta = \frac{\text{cov}(W_2, \ln(InExp))/\text{var}(W_2)}{\text{cov}(W_1, \ln(InExp))/\text{var}(W_1)}$ is conceptualized to be the degree of selection on unobserved confounders relative to the degree of selection on observed variables. The second parameter R_{max} is the maximum amount of variation in the outcome variable that could be explained when all possible confounding variables are controlled for. $R_{max} = 1$ when we assume $\text{var}(v|InExp, \boldsymbol{W}_1, W_2) = 0$. For given values of δ, R_{max}, and the estimation model, Oster (2019) established the relationship so that the coefficient adjusted for such a level of confounding factors $\tilde{\beta}$ can be calculated. Alternatively, fixing R_{max} and $\tilde{\delta} = 0$, I can calculate the corresponding value of δ, which is presented in Table 10.