

# Examining the Relationship of Expenditures and Student Outcomes at Public 4-Year Institutions

*Duncan Clark and Leonard Wainstein*

*March 23, 2018*

## Introduction

Higher education institutions serve a public good; they educate their students with the hope that in the future these graduates will go on and contribute to society as a whole. However, the students also benefit, as they tend to be more successful than their non-university educated peers. Thus, since much of higher education is funded by the public purse, it becomes a matter for the public to decide how much to subsidize individuals' educations for the good of the whole, and in particular it becomes crucial that the public money is used efficiently. It is this efficiency that we concerned ourselves with in this project.

In this project we hoped to analyze quantitatively the relationship between student outcomes at public 4-year higher education institutions and those institutions' expenditures. Specifically we looked at 6-year graduation rates for first-time, full-time, bachelor's degree-seeking student cohorts and the number of bachelor's degrees awarded per full time equivalent student (FTE). We decided to look at two outcomes because certain institutions may be more concerned with its bachelor's degrees per FTE than its graduation rate (or vice versa). A small school with mostly full-time student cohorts would likely be more interested in increasing its graduation rate. However, a large school with many part-time students would also want to know if the rate at which it is producing graduates is keeping up with the total number of credits students are taking that institution.

We were interested in using methods developed in this course to predict the two outcomes, but we were also interested in making valid interpretations of our results. With this in mind, we first tried only including variables that we believed to be important as in Tandberg and Hillman (2014) and Research for Action's 2017 study Slaughter et al. (2017). We then tried carrying out automatic variable selection through Lasso regression and using these variables as the basis for our models.

The data we analyzed was obtained from the Delta Cost Project at American Institutes for Research, which provides a dataset for all higher-education institutions in the US for academic years 1987 through 2015 (academic years are indexed by the year during which conclude - e.g. the 2015 academic year is the 2014-15 school year). The data is publically available at: <https://www.deltacostproject.org/>.

## Data Processing

We restricted our sample to 4-year, public, primarily bachelor's degree-granting institutions in the 50 states (plus DC), and excluded any institutions that did not exist in all years from 2003 to 2015. Years from 2003 onwards generally had a low level of missingness in the outcomes of interest, as well as covariates regarding expenditures, so we chose this as our cutoff for inclusion. As for examining other types of institutions, this would have been interesting, but these institutions generally serve purposes other than offering bachelor's degree programs. So, to avoid issues of heterogeneity in our sample, we excluded them from our analyses. We also dropped schools that were missing values for one of the outcomes in any of the years from 2003 to

2015, and schools with missing values for a few other key variables in all of those years of data. Starting with 553 schools in 2003 to 2015, thinning the sample in the described way went as follows:

- Kept schools in the 50 states plus DC (544 institutions kept)
- Dropped schools that didn't exist in the entirety of the time range, academic years 2003-2015 (505 institutions kept)
- Dropped schools that were missing bachelor's degrees awarded or graduation rate in any of the years in the data (457 institutions kept)
- Dropped schools that gave out 0 bachelor's degrees in a year or awarded fewer bachelor's degrees than associate degrees in a year (445 schools kept)
- Dropped schools that were missing values in all years from 2003 to 2015 for a few key variables (432 schools kept)

Dollar values were scaled to 2015 dollars using CPI, to allow for fair comparison, and monetary totals were scaled by total full time equivalent enrollment to account for the size of the institutions. There are alternate inflation indices that may have been appropriate (for example, HEPI - higher education price inflation, which is specific to the higher education sector). However, these were not considered.

We also added state and year specific GDP and unemployment data in line with the approach in Tandberg and Hillman (2014). The state GDP data came from the St. Louis Federal Reserve Bank's website for economic research (<https://fred.stlouisfed.org/>) and the state unemployment data came from the U.S. Department of Labor, Bureau of Labor Statistics website (<https://www.bls.gov/>).

## Training and Testing Split

To validate our models, we split our data into a test set and a training set. The test set was utilized only after models had been generated on the training set and it did not inform our modeling decisions (e.g., such as dropping of variables).

Since each institution appears in our dataset 13 times (once for each year in 2003-2015), we clustered sampled by institution in splitting our dataset, selecting all years of data into the test or training set for each institution. In addition, we stratified by census division, to assure that each region would be represented in our training and testing set and to reduce the variance in our estimators, as Tandberg and Hillman (2014) identified location as an important predictor of bachelor's degree-related student outcomes. The census divisions were defined as below:

*Table 1: States in each Census Division*

Region	States
New England	CT, ME, MA, NH, RI, VT
Middle Atlantic	NJ, NY, PA
East North Central	IN, IL, MI, OH, WI
West North Central	IA, KS, MN, MO, NE, ND, SD
South Atlantic	DE, DC, FL, GA, MD, NC, SC, VA, WV
East South Central	AL, KY, MS, TN
West South Central	AR, LA, OK, TX
Mountain	AZ, CO, ID, NM, MT, UT, NV, WY
Pacific	AK, CA, HI, OR, WA

We decided against including state fixed effects in any of our models, as we worried state fixed effects may account for too much variation, leading to poor prediction. We also decided against stratifying by state when splitting our data into a training and test set, and there were some states with only one institution in them, meaning states would be unrepresented in the training or test set.

After splitting the data, our training set and test set looked as follows:

*Table 2: Training and Testing Split*

	<b># of Institutions</b>	<b># of Obs.</b>
Training	347	4511
Testing	85	1105
Total	432	5616

## Models Including Variables per Literature

### Set-up

In attempt to hone in on the effect that changes in institution spending strategies have on student outcomes, we focused on the nine types of expenditure variables that were available to us in all institutions and years. These expenditure types are:

*Table 3: Expenditure Types and Examples*

<b>Expenditure Type</b>	<b>Example(s)</b>
Instruction	Teacher Salaries
Public Service	Conferences, Community Services
Academic Support	Libraries, Museums, Demonstration Schools
Student Services	Student Activities/Organizations
Institutional Support	General Administration
Operation and Maintenance	Utilities
Depreciation	Cost of Capital Assets
Scholarships and Fellowships	Scholarships, Fellowships
Auxiliary Enterprises	Residence Halls, Student Health Services, Intercollegiate Sports

Along with these expenditure variables, which were converted to 2015 values and divided by FTE enrollment, we also included several institution and state characteristics as per the two research studies mentioned before. These extra covariates are listed on the next page in Table 4.

Table 4: Covariates Included as per Literature

<p style="text-align: center;"><b><u>State-level Variables</u></b></p> <p style="text-align: center;">GDP per Capita (state and year-specific) Unemployment Rate (state and year-specific)</p>
<p style="text-align: center;"><b><u>Institutional Characteristics</u></b></p> <p style="text-align: center;">Undergraduate Enrollment % of Total Enrollment that is Under-represented Minority (Black or Hispanic) Revenue from Pell Grants per FTE Revenue from State Appropriations per FTE Net Tuition and Fees Revenue per FTE Total Revenue per FTE</p>
<p style="text-align: center;"><b><u>Categorical Variables</u></b></p> <p style="text-align: center;">Year Dummies (2003-2015) Census Division Dummies (9 Divisions)</p>

We performed minor missing imputation on the expenditure variables and above covariates by imputing with the predicted values of a linear regression model including only dummies for institution and year.

## Models

With these self-selected variables, we tried seven types of models on each of the two outcome variables:

- Traditional Linear Regression (LM)
- Traditional Linear Regression with Interactions
- Lasso Regression
- Lasso Regression with Interactions
- Ridge Regression
- Ridge Regression with Interactions
- Gaussian Kernel Regression

For the models with interactions, we interacted the nine expenditure variables with each other, as we hypothesized that spending more money in one expenditure area may change the effects that spending more money in another expenditure area have on our two outcomes (e.g., spending more money on instruction may have a stronger effect if the institution has already invested heavily in making students feel comfortable on campus, such as high spending for residence halls or student health services).

For the Lasso and Ridge models, we performed ten-fold cross-validation on the training set to estimate the regularization parameter ( $\lambda$ ). For the Gaussian Kernel model, we chose the  $\lambda$  that minimized the leave-one-out cross-validation error on the whole training set.

The testing root mean squared error (RMSE) of the above models on both outcomes is provided in Table 5 below:

Table 5: Testing Raw RMSE of Self-Selected Models (standardized RMSE in parentheses)

Model	6-Year Graduation Rate	Bachelor's Degrees per FTE
LM	0.1008 (0.6385)	0.0348 (0.9194)
LM (with interactions)	0.0983 (0.6230)	0.0359 (0.9461)
Lasso	0.1004 (0.6359)	0.0346 (0.9129)
Lasso (with interactions)	0.0979 (0.6205)	0.0337 (0.8895)
Ridge	0.1005 (0.6366)	0.0344 (0.9074)
Ridge (with interactions)	0.0990 (0.6276)	0.0331 (0.8734)
Gaussian Kernel	0.0959 (0.6076)	0.0274 (0.7235)

For 6-year graduation rate, the gaussian kernel model had the lowest RMSE. However, the kernel model performed only slightly better than the Lasso model with interactions (the model with the second lowest testing RMSE), and in general, the linear models performed quite well on the test data compared to the kernel model. Among the linear models, the models with the interaction terms performed better than the models without interactions, lending support to our hypothesis that there is some interaction effect between the expenditure variables.

For bachelor's degrees per FTE, the gaussian kernel model again had the lowest RMSE, and in this case performed much better than the linear models. And again, the linear models with interaction effects generally performed better than those without them. However, looking at the standardized errors in the above table (in parentheses), the models for bachelor's degrees per FTE performed much worse at predicting the outcome than those for 6-year graduation rate.

## Discussion

Because the kernel models performed the best on the testing set in terms of RMSE, we decided to interpret those models. In Table 6 below are the expenditure average partial derivatives from the kernel model that were significant at the 0.05 level for the 6-year graduation rate model.

Table 6: Significant (at the 0.05 level) Average Partial Derivatives of Expenditure Variables for the 6-Year Graduation Rate Kernel Model

Expenditure	Estimate	Standard Error	p-value
Academic Support Expenditures per FTE	4.50e-06	1.20e-06	0.0001495
Student Services Expenditures per FTE	1.62e-05	1.70e-06	0
Institutional Support Expenditures per FTE	-1.25e-05	1.20e-06	0
Depreciation Expenditures per FTE	2.90e-06	7.00e-07	0.0000419
Auxiliary Enterprises Expenditures per FTE	2.07e-05	6.00e-07	0

For 6-year graduation rates, student services and auxiliary enterprises expenditures had the most positive average partial derivatives and instruction expenditures did not end up being significant at the 0.05 level.

While we are wary of declaring any causality from these models, given that all of the expenditure variables were included in these models, as well as the total revenue per FTE that the school has at its disposal in the year (Total Revenue per FTE), this is interesting finding. Furthermore, this finding is consistent with the average partial derivatives from the bachelor's per FTE model, shown in Table 7 below.

*Table 7: Significant (at the 0.05 level) Average Partial Derivatives of Expenditure Variables for the Bachelor's Degrees per FTE Kernel Model*

<b>Expenditure</b>	<b>Estimate</b>	<b>Standard Error</b>	<b>p-value</b>
Public Service Expenditures per FTE	-7.00E-07	3.00E-07	0.0220192
Student Services Expenditures per FTE	4.50E-06	6.00E-07	0
Operation and Maintenance Expenditures per FTE	-1.40E-06	6.00E-07	0.0124906
Depreciation Expenditures per FTE	-6.00E-07	3.00E-07	0.0353355
Scholarships and Fellowships Expenditures per FTE	1.40E-06	6.00E-07	0.0162623
Auxiliary Enterprises Expenditures per FTE	2.40E-06	2.00E-07	0

It would be irresponsible to suggest that money spent to improve instruction does not have an effect on graduation rates or bachelor's degrees awarded per FTE (let alone the quality of education for all those at the school), and that all extra dollars should be funneled into student activities, residence halls, intercollegiate sports, or health services. However, these models do suggest that efforts to improve students' college experience and help students feel comfortable and supported on campus are not without benefits to a school's student outcomes.

As for the negative partial derivatives from these two models, namely institutional support for graduation rate and public service, operation and maintenance, and depreciation for bachelor's degrees per FTE, it would be easy to construct stories to explain these findings - perhaps schools that spend more on administration (which falls into institutional support) are more disorganized, or schools that spend more on maintaining their campus (operation and maintenance) are too concerned with cosmetics.

However, like for any positive partial derivatives from the above, it is important to keep several things in mind. Firstly, these are *average* partial derivatives from the training sample and the partial derivative for a school for one of these expenditures depends on the school's values for all of the variables that were included in these models. In Table 8 and Table 9 below are, for the expenditures with the most positive average partial derivatives, the percent of partial derivatives that were less than zero, and, for the expenditures with the most negative average partial derivatives, the percent of partial derivatives that were greater than zero.

*Table 8: % of Partial Derivatives Less than 0 for Most Positive Average Partial Derivatives*

<b><u>Graduation Rate</u></b>	<b><u>% of Partial Derivatives Less than 0</u></b>
Student Services Expenditures per FTE	18.2%
Auxiliary Enterprises Expenditures per FTE	0.9%
<b><u>Bachelor's per FTE</u></b>	<b><u>% of Partial Derivatives Less than 0</u></b>
Student Services Expenditures per FTE	23.4%
Auxiliary Enterprises Expenditures per FTE	4.8%

Table 9: % of Partial Derivatives Greater than 0 for Most Negative Average Partial Derivatives

<u>Graduation Rate</u>	<u>% of Partial Derivatives Greater than 0</u>
Institutional Support Expenditures per FTE	3.2%
<u>Bachelor's per FTE</u>	<u>% of Partial Derivatives Greater than 0</u>
Public Service Expenditures per FTE	19.1%
Operation and Maintenance Expenditures per FTE	26.9%

Even the expenditures with the most negative average partial derivatives for the bachelor's degrees per FTE model had nearly over 20% of partial derivatives greater than 0 in the training set. And for both outcomes, the percentage of partial derivatives less than zero for student services was also around 20% of the training set. So, these relationships are not so uniformly positive or negative.

The second thing to keep in mind is effect size. Assuming causality, if a school were to transfer a large portion of its funds to auxiliary enterprises (which had the most positive average partial derivative in both models) from the expenditure area with the most negative partial derivative, what would be the effect on these outcomes be? Running first differences by increasing auxiliary enterprises expenditures per FTE by 100 and subtracting 100 from the expenditure per FTE variable with the most negative average partial derivative for each outcome (in many cases, a drastic and unrealistic change in a school's spending strategy) yields the following average changes in the outcomes:

Table 10: First Differences of increasing Auxiliary Enterprises Expenditures per FTE by 100 and Subtracting 100 from the Expenditure per FTE with Most Negative Average Partial Derivative from each Model

<b>Outcome</b>	<b>First Difference Estimate</b>
Graduation Rate	+0.00326
Bachelor's Degrees per FTE	+0.0006

For both outcomes, the effect is extremely small for such a drastic change in the spending strategy of a school - on average, 0.3% more students in a cohort graduated for the graduation rate first difference, and 0.0006 more bachelor's degrees granted per FTE in a year.

Finally, it is important to keep misspecification in mind. Are the variables included in these models sufficient to understand the relationship between a school's expenditures and student outcomes? As we will see in the next sections, perhaps not. By viewing this problem through a more predictive lens, we were able to find models that greatly reduced the testing error for each outcome.

## Automatic Variable Selection

### Setup

For this approach, the idea was to leave as much information available to the model as possible. Variables that were obviously colinear and not of interest were excluded, e.g. full time retention rate is a strong predictor of

graduation rate. However, this is not of interest, since it seems obvious that if students remain in full time enrollment, they will graduate within a specified time.

With around 4000 observations, variables with more than 200 missing observations were dropped. The remaining missing values were imputed with a linear regression model including only dummies for region and year.

## Selection

Since Lasso induces sparsity in the coefficient estimates, i.e it sets many of them to be 0, we used it to automatically select the variables with high predictive power for both our outcomes. We ran 10-fold cross validation to select the optimal levels for the regularization parameters (10 here is arbitrary, though does not seem unreasonable since it is small in relation to our ~4000 observations).

We selected the 10 variables with the largest absolute value coefficients. Again, the 10 here is arbitrary, though we suspect our models are robust to this since the size of the coefficients does decrease rapidly due to the sparsity-inducing Lasso.

Table 11 and 12 give the selected variables for each outcome as well as the value of their Lasso coefficients.

The full definition of variables is given in the data dictionary publically available at <https://www.deltacostproject.org/>.

There are three main types of variable that the Lasso model has identified as strong predictors for both graduation rate and numbers of bachelors degrees per fte:

- part time variables
- expenditure variables
- grant variables

Part time variables were selected multiple times in both models, suggesting that there is an association between schools with higher portions of part time students and poorer student outcomes. Because graduation rates only account for full time students, we hypothesise that this link is identified due to institutions that have a higher proportion of part-timers being more amenable to students dropping to part time study, which would likely lower graduation rates. However, interestingly, they are also selected for the bachelors per FTE, which does account for part time students. This may suggest that part time students are less likely to graduate even when accounting for their reduced hours.

The selection of the expenditure variables is promising and supports the hypothesis that an institution's expenditures are related to its student outcomes.

We included grants in our first models due to their inclusion in Tandberg and Hillman (2014) and Slaughter et al. (2017), which use the grant variables as a proxy for deprivation of the student body, i.e. students from a lower socio - economic group are more likely to be awarded a federal pell grant. The inclusion of these variables for both outcomes in the automatic selection suggests they are good predictors of outcomes and suggests that including them in our first models did serve to reduce the variance of our estimators.



The signs of the coefficients for graduation rate are as expected, with high levels of part time study and federal grants having a negative impact and spending having a positive impact.

The coefficients for the Bachelor's degrees per FTE models are somewhat less interpretable. However, the majority of the coefficients do follow the same signs as for graduation rates, with the notable exception of `eandr_degree` - educational spending per degree. We would expect this to have a positive impact, not a negative impact. We think this is possibly due to high levels of correlation between one of the the selected variables and education spending, as well as the misspecification problems of the linear model. See below for our kernelized approach which helps to mitigate this problem.

## Models

We used classical, ridge, lasso and elastic net regression on the full data with all variables, and carried out kernel regularised regression with a Gaussian kernel on the variables selected by Lasso, for both outcomes. The results are summarised in Tables 13 and 14 below:

As expected, the root mean square error for each of the models is lower than those achieved using the variables suggested in the literature. Using all the variables, the Lasso performed best out of the regularised linear regression models for both outcomes.

We also note that the kernelized regression model performs the best overall for both outcomes, perhaps suggesting a complex non-linear relationship between the selected variables.

Adding the interaction terms for the Lasso selected variables, with the hope of obtaining non trivial first difference estimates, essentially just resulted in overfitting. We could have mitigated this by adding another Lasso or ridge penalty term.

As in the first method, we analyzed the Kernel models' average partial derivatives in an attempt to understand the mariginal effects of changes in the variables.

In general, Bachelor's per FTE was a harder variable to predict, with all of the models performing less well and as such our interpretation on this variable holds less weight. This is consistent with the findings using the self-selected variables.

## Discussion

Again, because the kernel models performed the best on the testing set in terms of RMSE, we decided to interpret those models. In Table 15 below are the average partial derivatives from the kernel model that were significant at the 0.05 level for the 6-year graduation rate model.

*Table 15: Top 10 Average Partial Derivatives of Lassoed Variables for the 6-Year Graduation Rate Kernel Model*

Variable	Estimate ( Standardised)	Standard Error	p-value
other full time	0.3928	0.011	0.0000
fed grant pct	-0.2004	0.0094	0.0000
grant01	-0.2792	0.011	0.0000
eandg02	0.1762	0.0099	0.0000
auxillary03	0.1442	0.01	0.0000
ptug share of total pt enrl	-0.1278	0.0074	0.0000
bach deg share of tot deg	0.1220	0.0077	0.0000
fall cohort pct	0.1073	0.0081	0.0000
total part time postbacc	-0.0625	0.0107	0.0000
total part time	-0.0966	0.0099	0.0000

Here it is noticeable out of the selected variables that the part time coefficients seem to have smaller effect sizes, with the exception of “other full time”, which has the largest average partial derivative. In this model, a small change in the “other full time” variable seems to account for the effect of a change in the part time proportion of an institution, since these variables are highly correlated in the data.

Federal grants have high negative average partial derivatives - agreeing with the Lasso model, our intuition, and our inclusion of federal grants in our self selected models.

Interestingly, like when we included all expenditure variables in our self-selected models, auxiliary revenue, which is highly correlated with auxiliary expenditures, had positive average partial derivative. However, with this model, education and related expenditures on salaries has a larger average partial derivative. We do not over-interpret the fact that salaries were selected by the Lasso since this is highly correlated with the total expenditures variable.

**Whereas our first approach identified this variable as having the largest average partial derivative. Thus there is some evidence that the tempting interpretation that auxillary expenditure is most important may be a model dependent and a step to far. However we do note that the fact that it was identified in both approaches as important does suggest that it has an effect. This in itself is interesting as it may be counterintuitive that spending not directly on education has an effect on the effectiveness of the institution as a whole, and perhaps justifies public spending on such activities, that could be deemed non essential when allocating public funds.**

The average partial derivatives for the variables in the bachelor’s degrees per FTE kernel model, provided in Table 16 below, align less well with our intuition and the first approach.

*Table 16: Top 10 Average Partial Derivatives of Lassoed Variables for the Bachelors per FTE Kernel Model*

Variable	Estimate (Standardised)	Standard Error	p-value
hbcu2	0.2078	0.0400	0.0000
other ed related cost	0.0971	0.0027	0.0000
fed grant num	0.0836	0.0029	0.0000
inst grant num	0.0353	0.0023	0.0000
credhoursug	0.0291	0.0028	0.0000
eandr degree	-0.0282	0.0034	0.0000
returning to total undergraduate	0.0191	0.0025	0.0000
ptug share of total pt enrl	0.0106	0.0022	0.0000
total enrollment black tot	-0.0029	0.0036	0.4119

However, we do note that part time effects seem to have diminished, having low average partial derivatives. This is somewhat expected, since the bachelor’s degrees per FTE outcome takes into account the level of part time studies at an institution. However, looking at the Lasso coefficients, we might have inferred an association with part timers and poorer outcomes, despite the outcome accounting for part time study. However, this seems not to be the case in light of the low coefficients in the kernel model for the part time variables.

Further interpretation is not obvious, though we note that the negative average partial derivative for eand degree is still present, which may suggest that the reason for the negative lasso coefficient was not due to model misspecification, as the kernel model is much more flexible. Overall, note that the effect size is much smaller here than it was for graduation rate, which matches with the bachelor’s degrees per FTE being a noisier outcome variable.

## Conclusions

We only feel comfortable making conclusions in terms of interesting associations, and would suggest more detailed research into the relationship of expenditure areas and student outcomes. However, our models suggest the following:

- Auxiliary services spending may worthwhile, as it is was found to be positively associated with 6-year graduation rates in both the self-selected models and the models that were developed using automatic variable selection.
- Because auxiliary services expenditures was the only expenditure variable for which the two modeling approaches matched up, we would refrain from making any conclusions or suggestions about the relationship between the two outcomes and another expenditure. While the models from the automatic approach have much more predictive power than the self-selected models do, we worry that approaching the problem purely from a predictive standpoint gets in the way of discovering the true relationship between the outcomes and the expense variables. And while the self-selected models would help in understanding this true relationship under correct specification, misspecification is certain.
- Bachelor’s degrees per FTE is much more difficult to predict than graduation rate. We hypothesize this is because unlike 6-year graduation rates, the bachelor’s degrees per FTE outcome takes into account part-time students, whose college careers are much less predictable than full-time students.

## Limitations

- All institutions were treated equally with no weighting based on size (e.g., UCLA has same effect as small college granting <100 degrees per year)
- With high levels of multicollinearity in expenditure variables, using the lasso for variable selection may not be much better than random selection

## References

- “Delta Cost Project.” 2018. Accessed February 15. <https://www.deltacostproject.org/>.
- Slaughter, A., K. Callahan, D. Kim, K. Meehan, L. Wainstein, L. Liu, and X. Zhu. 2017. “Playing the Long

Game: The Impact of Outcomes-Based Funding on Bachelor's Degree Attainment." Research for Action.

"St. Louis Federal Reserve Bank." 2018. Accessed February 14. <https://fred.stlouisfed.org/>.

Tandberg, and Hillman. 2014. "State Higher Education Performance Funding: Data, Outcomes, and Policy Implications." *Journal of Education Finance* 39 (3). University of Illinois Press: 222–43.

"U.S. Department of Labor, Bureau of Labor Statistics." 2018. Accessed February 14. <https://www.bls.gov/>.