Student Debt as Investment: ROI Analysis Across Institution Types

S470/670 · Fall 2025

Yash Takte & Meryl Sarah Jacob

1. Introduction and Statement of Goals

This study addresses the main question faced by prospective college students and policy makers: **How does student debt as an investment affect earnings, and does this ROI relationship vary across institution types (public or private nonprofit) and institutional quality (graduation rates)?**

We focus on two institutional dimensions that prior research and policies suggest are critical factors that influence debt. First, is 'Institution Type: Public vs Private nonprofit'. These represent the different funding models, cost structures and student population. These two different types of institutions may produce distinct debt-earning patterns. Second, is 'Institutional Quality'. This is captured through 6-year graduation rates which measure an institution's effectiveness at helping students complete degrees. Degree completion is traditionally the main gateway to labor market returns.

Understanding the debt-earnings relationship helps prospective students and families make consequential financial decisions. Conventionally, the advice is to minimize debts at all costs, potentially steering students towards lower-cost institutions with weaker completion rates and employment outcomes. If our analysis reveals that moderate debt at high-graduation-rate schools produces better outcomes than minimal debt at struggling schools, this advice may be misguided. Conversely, if we find danger zones where borrowing doesn't translate to earnings gains, students can avoid these combinations.

For policy makers, federal student loan policy currently treats all bachelor's degree programs similarly, i.e. undergraduates can borrow up to $31,000 in federal Stafford loans regardless of institutional quality, program of study or expected outcomes. If the debt-earnings relationship varies by institution type and quality, a one-size-fits-all policy may be inefficient. These policies could be then tailored to institutional characteristics and outcomes.

2. Data Description

2.1 Data Source and Variables

Data Source: The U.S. Department of Education, College Scorecard (https://collegescorecard.ed.gov/data/)

Data: https://ed-public-download.scorecard.network/downloads/Most-Recent-Cohorts-Institution_05192025.zip (Data retrieved December 9, 2025)

We use data from the Official Federal database of college performance and student outcomes which includes the most recent Institution-Level Data (this file contained information about debt and earnings). To this data we applied several filters to narrow it down for our problem statement. We finally chose 1,430 bachelor's degree-granting institutions (filtered from 6,429 total). These represented most U.S. four-year colleges. 1,430 institutions are distributed as Private Nonprofit (889 or 62.2%), Public (507 or 35.5%), and For-Profit (34 or 2.4%). We also narrow down the control variables and drop the rest.

Table 1: Data Cleaning Summary

Data Cleaning Step	Number of Institutions
Original dataset	6,429
After filtering (4-year only)	3,039
After removing missing data	2,400
Final complete case dataset	1,430

Table 2: Filtering Criteria (Final Dataset)

	Criterion
1.	4-year institutions (ICLEVEL = 1)
2.	Predominantly bachelor's degree granting (PREDDEG = 3)
3.	Complete earnings and debt data
4.	Complete control variables (graduation rate, admission rate, % Pell)
5.	Mainland U.S. only (excludes territories)

Table 3: Distribution by Institution Type (Final Dataset)

Institution Type	Count	Percentage (%)
Public	507	35.5
Private Nonprofit	889	62.2
For-Profit	34	2.4

2.2 Variable Definitions

The key variables (control variables) are:

- Median Debt at Graduation - Median Federal loan debt at graduation
- Median Earnings (1 Year After) - Median Income one year after graduation
- Graduation Rate - Percentage of students completing their degree within 6 years
- Institution Type - Public, Private Nonprofit, or For-Profit

Table 4: Key Continuous Variable Ranges (Final Dataset)

Variable	Min	Max	Median
Earnings (1yr)	$14,067	$122,568	$43,259
Debt	$3,591	$42,125	$23,428
Graduation Rate	0.0%	100.0%	58.1%

2.3 Understanding Data Artifacts

Before analyzing relationships, we must understand borrowing constraints. About three-fourths of private nonprofit student debt clusters at the $25K-$27K federal Stafford loan limit. On the other hand, public schools show a lower debt of about $21K. This clustering indicates the federal policy creating a ceiling on borrowing for many dependent undergraduate students.

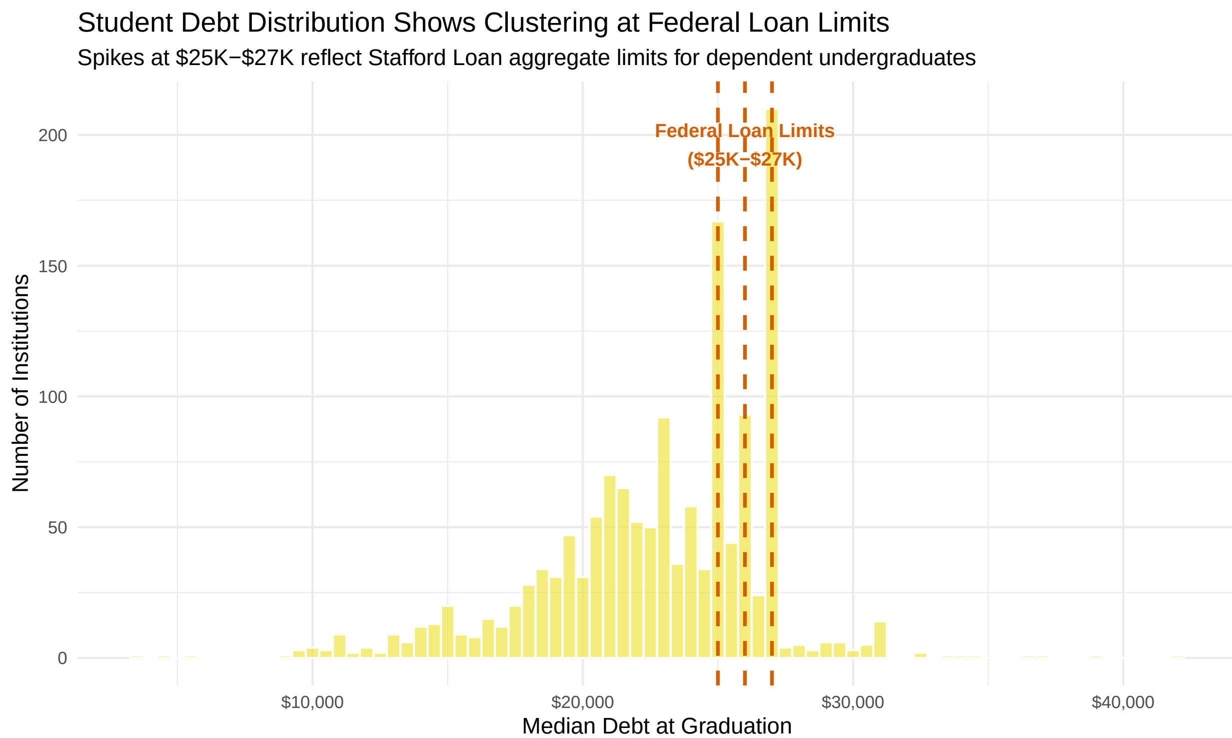
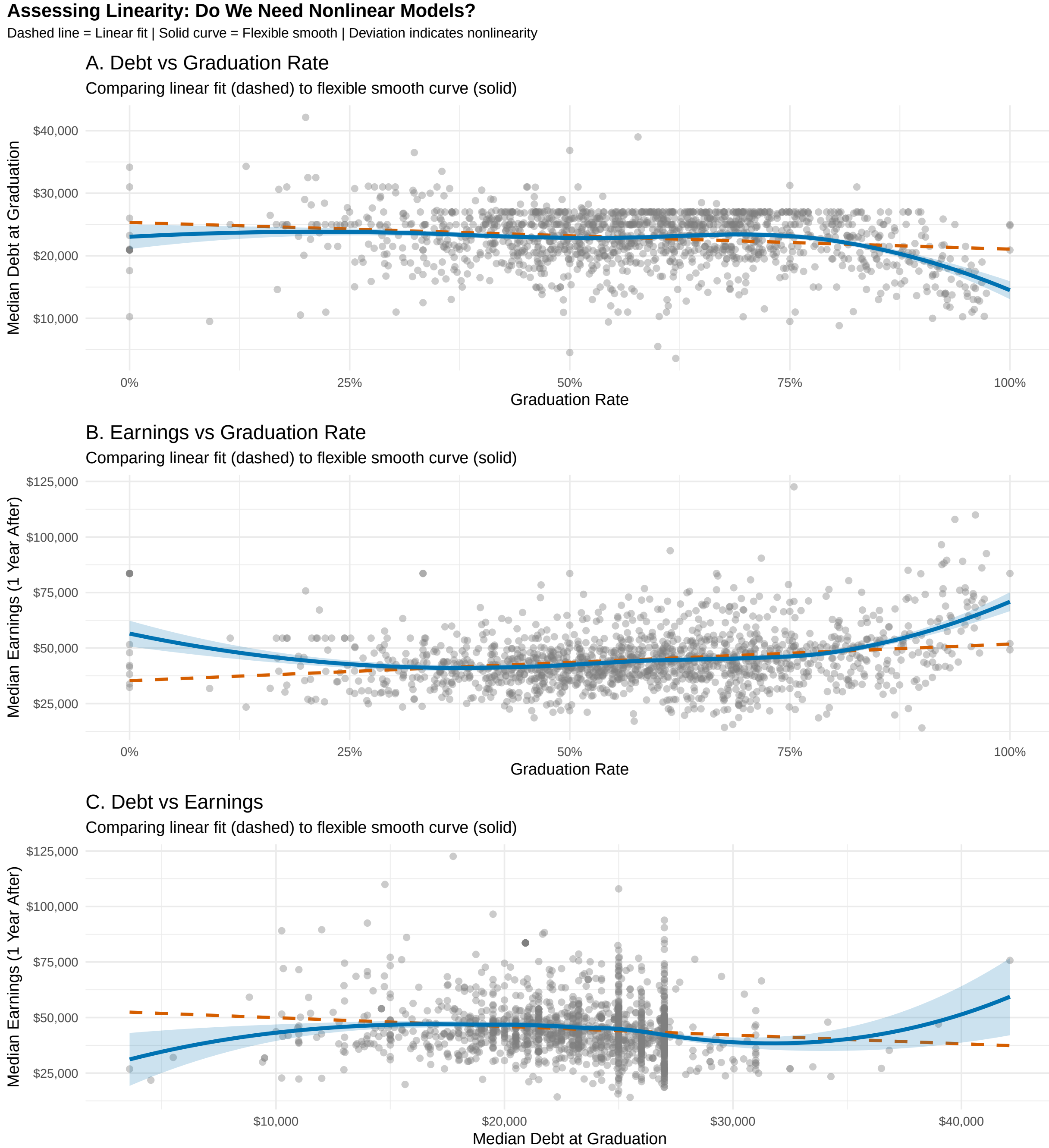


Table 5: Debt Clustering Patterns by Institution Type

control_label	N	Median Debt	% Between $22K-$32K	% Above $30K
Public	507	$21,056	38.9	1.8
Private Nonprofit	889	$25,000	76.2	1.6
For-Profit	34	$23,597	47.1	20.6

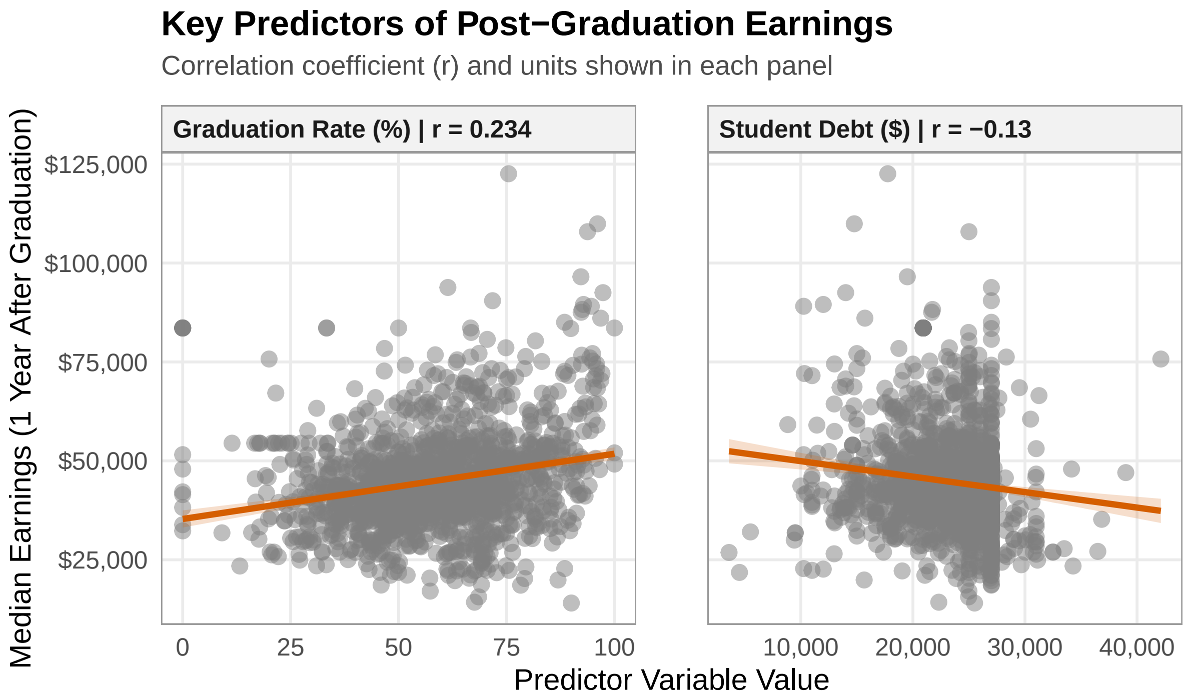
3. Exploratory Data Analysis

3.1 Assessing the Need for Nonlinear Models



The graphs above check to see if we need complex models, or will simple straight lines suffice. We compare linear fits (dashed) to flexible smooth curves (solid). In Panel B (Earnings vs Graduation Rate), the smooth curve reveals a steepening relationship; graduation rate matters more at higher levels. We see a steady increase in the relationship after a point. In Panel C (Debt vs Earnings), the smooth curve shows a subtle U-shape that a straight line misses entirely. These deviations indicate that linear models will miss important structure in the data.

3.2 Key Predictors of Earnings



Across all 1,430 institutions, graduation rate correlates at just r=0.234 with earnings, while debt shows a weakly negative r=-0.13. These anemic correlations might suggest neither factor has a large effect. But aggregating across all school types might obscure critical differences.

3.3 The Role of Graduation Rate

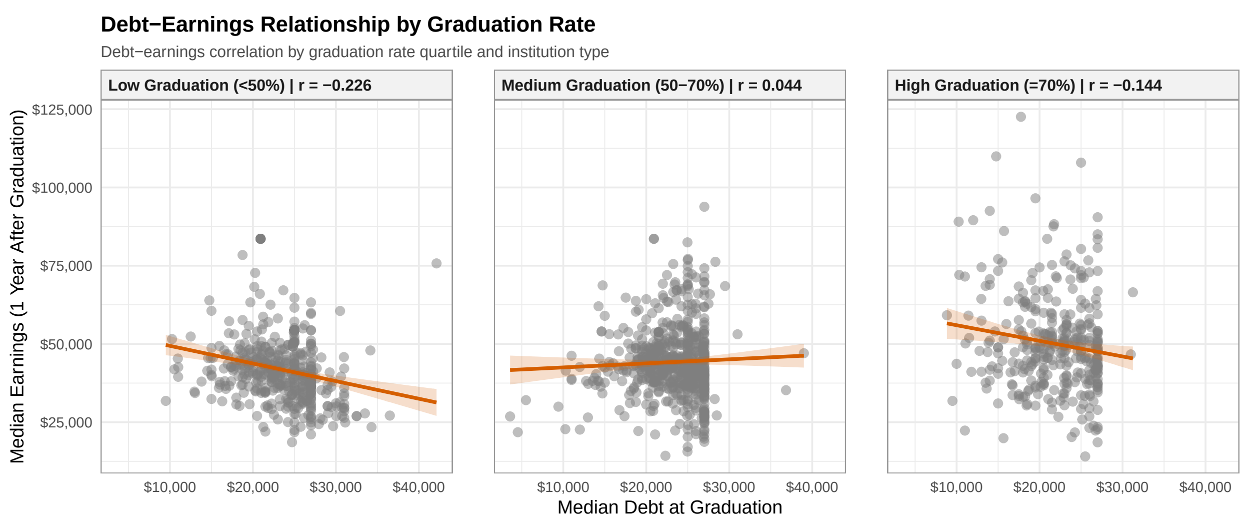


Table 6: Debt-Earnings Correlation by Graduation Rate Tier

grad_category	N	Correlation
Low Graduation (<50%)	461	-0.226
Medium Graduation (50-70%)	616	0.044
High Graduation (70%)	353	-0.144

To examine whether the debt–earnings relationship varies by institutional quality, we stratified institutions into three graduation-rate tiers (<50%, 50–70%, 70%) and fit separate linear models within each group. If debt had a uniform association with earnings, slopes would be similar across tiers. Instead, we observe substantial heterogeneity, including a reversal in the direction of the relationship across graduation tiers, indicating a clear interaction between debt and institutional quality.

This pattern reversal motivates our analytical approach: a single linear model assuming a constant debt effect is inappropriate. The data requires flexible models capable of capturing interaction and nonlinearity. Accordingly, we employ Generalized Additive Models (GAMs) to model these context-dependent relationships while limiting overfitting.

4. Model Development and Selection

We fit sector-specific GAMs using log-transformed earnings and debt to handle skewness and proportional relationships. We used scaled t-distributions (family=scat) for outlier resistance and REML for smoothing parameter selection. For each sector, we tested three specifications: (A) separate smooths for debt and graduation rate, (B) smooth debt with linear graduation rate, and (C) interaction surfaces.

4.1 GAM Model Variants

Table 7: Model Comparison: Which Option is Best?

Sector	Model	AIC	Dev. Explained	R^2	ΔAIC
Public	Option A: Both Smooth (Debt & Grad Rate)	-444.5	20.3	0.183	11.4
Public	Option B: Debt Smooth, Grad Linear	-369.4	11.1	0.106	86.5
Public	Option C: Interaction Surface	-455.9	22.2	0.194	0.0
Private Nonprofit	Option A: Both Smooth (Debt & Grad Rate)	276.0	14.2	0.142	0.0
Private Nonprofit	Option B: Debt Smooth, Grad Linear	304.2	10.6	0.108	28.1
Private Nonprofit	Option C: Interaction Surface	279.2	14.3	0.137	3.2

For public institutions, the interaction surface (Option C) won decisively, explaining 22.2% of deviance. This interaction model captures how debt's effect changes at different graduation rates, which is essential for public schools where the reversal pattern was strongest. For private nonprofits, separate smooths (Option A) performed slightly better than the interaction, suggesting additive rather than interactive effects. These model selections aren't arbitrary, but they reflect fundamentally different data-generating processes across sectors.

4.2 Addressing Overfitting Concerns

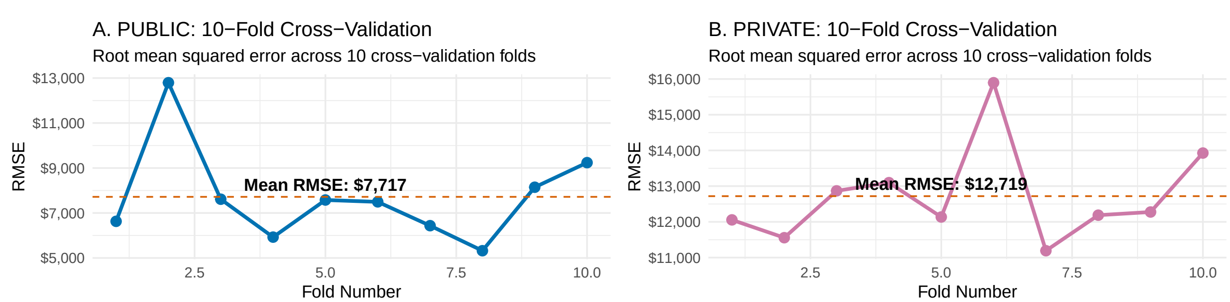


Table 8: Cross-Validation Results: Model Stability Across Folds

Sector	Mean RMSE	SD of RMSE	CV (%)
Public	$7,717	$2,111	27.4
Private Nonprofit	$12,719	$1,361	10.7

Because our models are relatively complex, overfitting was a concern. We evaluated generalization using 10-fold cross-validation, repeatedly training on 90% of the data and testing on the remaining 10%. The public-school model shows stable performance with an average RMSE of $7,717, while the private-school model has a higher average RMSE of $12,719. The low variability in error across folds indicates good generalization. The higher error for private institutions reflects greater sector heterogeneity rather than model instability.

5. Results

5.1 Model Validation

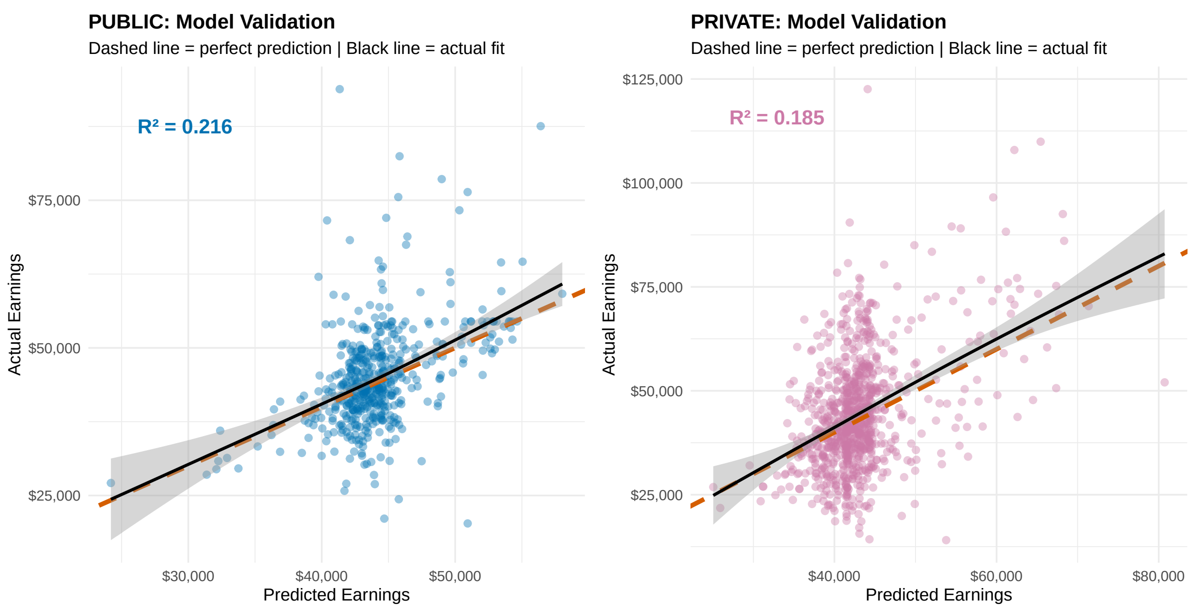
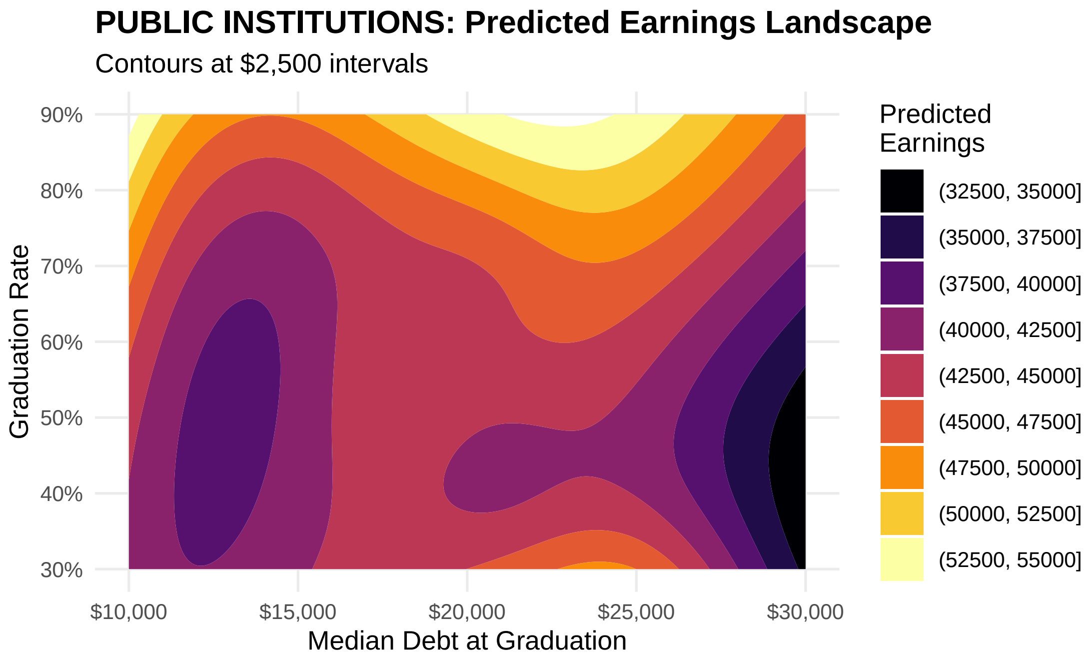


Table 9: Model Performance: Public vs Private Nonprofit GAMs

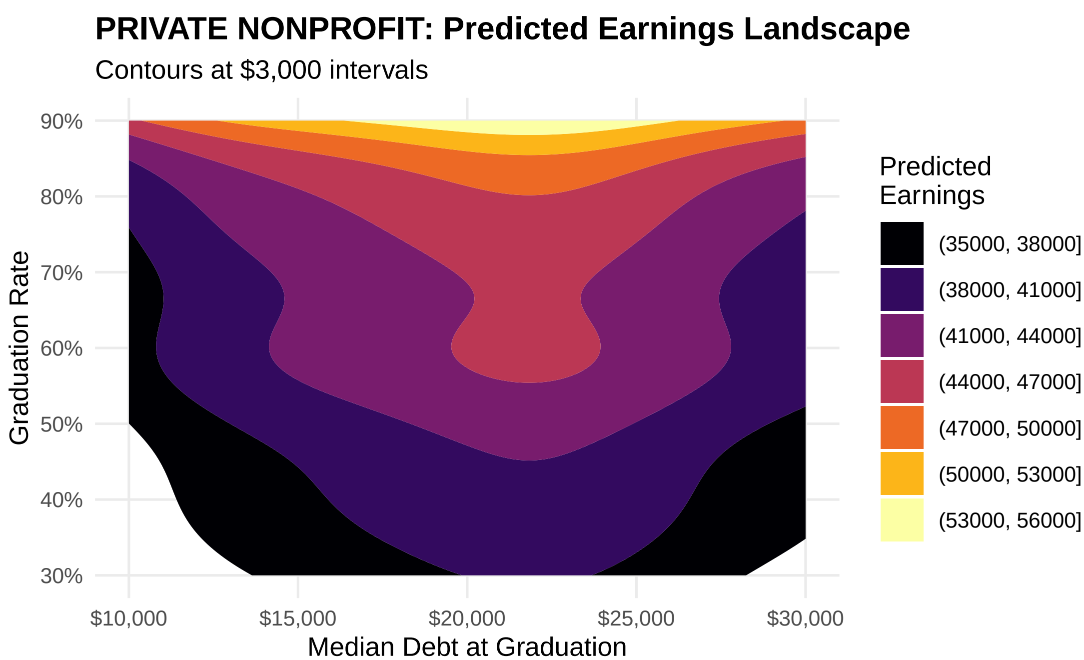
Metric	Public	Private
R^2 (Variance Explained)	0.216	0.185
RMSE (Avg $ Error)	$7,697	$12,529
MAE (Typical $ Error)	$5,079	$9,376
MAPE (Avg % Error)	11.4%	22.7%

Our public model achieves R^2=0.216 (explaining 21.6% of earnings variance) with typical errors around $5,079 (MAE). The private model achieves R^2=0.185 with MAE=$9,376. While these R^2 values might seem modest, they represent only one-year earnings predictions using just two institutional characteristics. This also suggests that a majority of the variance comes from factors we don't measure like field of study, geographic location, career path choices, etc. What the above model captures is the portion of earnings attributable to institutional debt levels and completion effectiveness.

5.2 The Earnings Landscape: Contour Plots



This contour map reveals the predicted earnings landscape for public universities and shows a vertical stratification (contour lines run nearly horizontal), meaning earnings change dramatically as you move up in graduation rate but modestly as you move right in debt. A student moving from a 40% to 80% graduation-rate public school gains approximately $15,000 in predicted earnings, even if debt increases from $15K to $25K. The highest earnings zone (bright yellow, $50K-$55K) occurs at high graduation rates with low to moderate debt. There's a 'valley' around 40-50% graduation rates where earnings dip regardless of debt; these mid-tier schools with low completion rates produce the weakest returns. The jagged contours and multiple local peaks reflect the interaction model capturing complex, non-monotonic patterns that simpler models would miss.

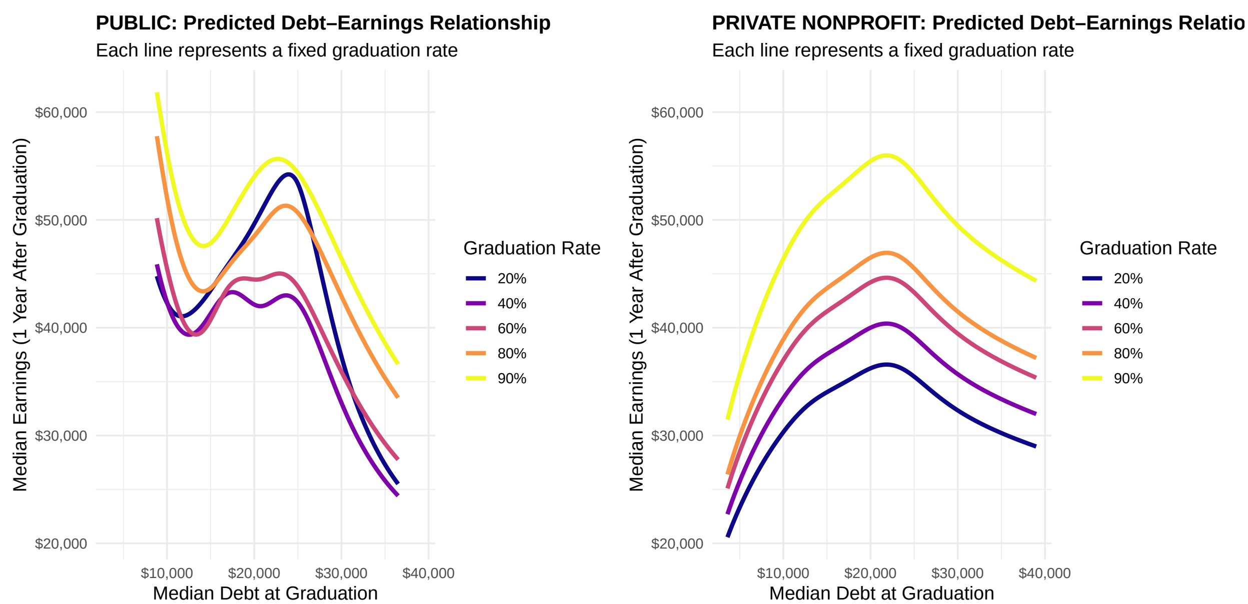


The private nonprofit landscape looks fundamentally different with smoother, more regular contours. Here we see horizontal gradients, where earnings increase substantially with higher debt even when graduation rate stays constant. A private school student with \$30K debt at a 60% graduation-rate school is predicted to earn around \$47K, while a peer with just \$15K debt at an equivalent-quality school earns closer to \$41K.

Higher debt at private schools likely signals attendance at more expensive, prestigious institutions with superior resources, networks, and brand name. Higher debt does not cause higher earnings; rather, debt could also correlate with institutional characteristics (selectivity, endowment, alumni networks) that we don't directly measure. The smoother contours reflect our additive model, i.e. debt and graduation rate each contribute independently rather than interacting.

5.3 Predicted Debt-Earnings Curves

To make these contour maps concrete, we trace how predicted earnings change with debt at fixed graduation rates.



For public schools, the curves diverge dramatically by graduation rate. The 90% line (yellow) starts around \$50K at low debt and stays relatively flat; these high-quality schools deliver strong outcomes regardless of moderate borrowing. The 40% line (purple) starts low (~\$40K) and rises before declining, suggesting students at struggling schools who can't afford any borrowing may not complete degrees, while those with \$20-25K debt can finish and see returns. However, beyond \$25K debt, returns collapse.

For private nonprofits, all curves follow similar inverted-U shapes, peaking around \$15-20K debt regardless of graduation rate. This consistency validates our additive model, i.e. debt operates similarly across quality tiers. But crucially, higher graduation rates shift the entire curve upward by \$5-10K, showing quality still matters even if the debt-earnings shape is consistent.

5.4 Which Universities Fit Well? Which Don't?

This section captures which schools our model captures well, and where does it fail. This addresses a key concern: are we just fitting noise, or capturing real patterns?

Table 10: PUBLIC: Best Model Fits (Top 5)

name	earnings	predicted	residual	grad_rate	debt
Emporia State University	$44,153	$44,143.25	$9.75	54.3%	$19,500
Indiana University-Bloomington	$48,621	$48,631.02	-$10.02	81.2%	$19,509
Louisiana Tech University	$44,484	$44,512.09	-$28.09	57.6%	$22,135
Western Illinois University	$41,243	$41,205.50	$37.50	44.8%	$25,251
Central Washington University	$43,326	$43,364.74	-$38.74	49.8%	$19,500

Table 11: PUBLIC: Worst Model Fits (Top 5)

name	earnings	predicted	residual	grad_rate	debt
Maine Maritime Academy	$93,824	$41,347.04	$52,476.96	61.4%	$27,000
California State University Maritime Academy	$82,458	$45,832.60	$36,625.40	66.8%	$24,965
South Dakota School of Mines and Technology	$71,587	$40,399.80	$31,187.20	58.2%	$27,000
Georgia Institute of Technology-Main Campus	$87,556	$56,408.39	$31,147.61	92.3%	$21,672
University of North Carolina School of the Arts	$20,265	$50,937.31	-$30,672.31	79.2%	$23,870

Table 12: PRIVATE: Best Model Fits (Top 5)

name	earnings	predicted	residual	grad_rate	debt
Central Methodist University-College of Liberal Arts and Sciences	$41,468	$41,452.47	$15.53	52.5%	$17,619
Manchester University	$38,243	$38,304.13	-$61.13	44.8%	$26,854
Cedarville University	$45,463	$45,363.29	$99.71	74.3%	$20,937
Regent University	$42,499	$42,615.59	-$116.59	54.1%	$24,534
University of Richmond	$52,666	$52,528.03	$137.97	87.8%	$21,000

Table 13: PRIVATE: Worst Model Fits (Top 5)

name	earnings	predicted	residual	grad_rate	debt
University of Health Sciences and Pharmacy in St. Louis	$122,568	$44,099.61	$78,468.39	75.5%	$17,755
Molloy College	$90,478	$41,888.45	$48,589.55	71.8%	$27,000
Harvey Mudd College	$107,923	$62,196.58	$45,726.42	93.8%	$25,000
Massachusetts Institute of Technology	$109,923	$65,437.53	$44,485.47	96.1%	$14,768
The Juilliard School	$14,067	$53,780.60	-$39,713.60	90.0%	$25,500

Our best fits include mainstream universities like Indiana University-Bloomington (public) and Cedarville University (private), schools where debt and graduation rate accurately predict earnings. These schools validate the fact that our models work for typical institutions.

Our worst fits show that we overpredict earnings at the University of North Carolina School of the Arts (predicted \$51K, actual \$20K). Arts programs produce graduates with low initial earnings despite high graduation rates. We underpredict at Maine Maritime Academy (predicted \$41K, actual \$94K) and Massachusetts Institute of Technology (predicted \$65K, actual \$110K). Maritime academies and elite STEM schools produce earnings our general model can't capture because program-specific factors (engineering, maritime careers) dominate.

These outliers aren't failures but reveal that field of study and program specialization matter enormously. Our model only captures what debt and completion rates can explain.

6. Conclusions

Analyzing 1,430 four-year U.S. institutions, we find that the relationship between student debt and earnings varies sharply by institutional context, undermining narratives that treat borrowing as uniformly beneficial or harmful.

Key Findings:

1) The debt–earnings relationship is nonlinear and context-dependent. The weak overall correlation masks substantial heterogeneity: when stratified by graduation rates, the association reverses across tiers, indicating that institutional effectiveness critically shapes outcomes.
2) Public and private nonprofit institutions exhibit distinct dynamics. At public institutions, graduation rates dominate earnings outcomes, with higher-quality schools delivering substantially higher earnings even at higher debt levels. In contrast, private nonprofits display flatter debt–earnings relationships, likely reflecting institutional prestige and unobserved resources rather than debt's causal impact.
3) We identify clear risk and opportunity zones. High borrowing at lower- and mid-graduation public institutions is associated with poor outcomes, while moderate debt at high-graduation public schools appears economically justifiable.

These findings reflect associations rather than causal effects. Selection bias and omitted variables, most notably field of study, likely influence observed patterns, particularly at specialized institutions, and should be addressed in future work.

7. Limitations

Unmeasured confounding factors dominate: Our models explain only 18-22% of earnings variation. The remaining 78-82% reflects unmeasured factors: field of study (engineering vs humanities differs by \$30,000+ annually), student ability, family wealth, geography, and career choices.

Data constraints limit interpretation: We measure only one-year post-graduate earnings for completers, missing long-term trajectories and non-completers (often worst outcomes). We lack program-level data, forcing institution-level analysis that obscures within-school variation.

Causality cannot be established: Students self-select into institutions and debt levels based on unmeasured characteristics. Higher debt at private schools likely reflects selection into prestigious institutions rather than debt's causal effect on earnings.

8. Future Scope

Further analysis should account for field of study, as debt–earnings relationships may differ substantially across disciplines and may explain much of the observed institutional variation. Extending the analysis to longer-term earnings would clarify whether short-term patterns persist or change as loan repayment burdens accumulate. Including non-completer data is critical to capture worst-case outcomes for borrowers who incur debt without earning a credential. Finally, incorporating geographic cost-of-living adjustments would distinguish nominal earnings differences from real economic outcomes.

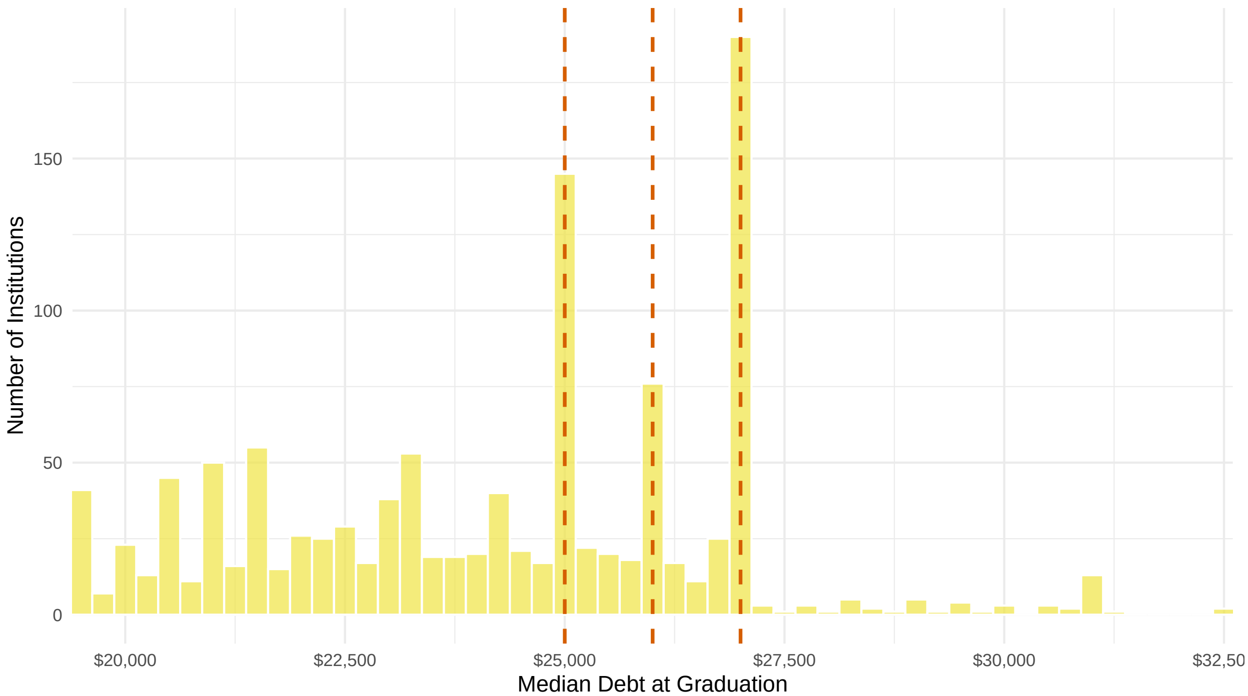
References

U.S. Department of Education, Federal Student Aid. (n.d.). *Loan Amount Limits.* Retrieved December 9, 2025, from https://studentaid.gov/understand-aid/types/loans/subsidized-unsubsidized

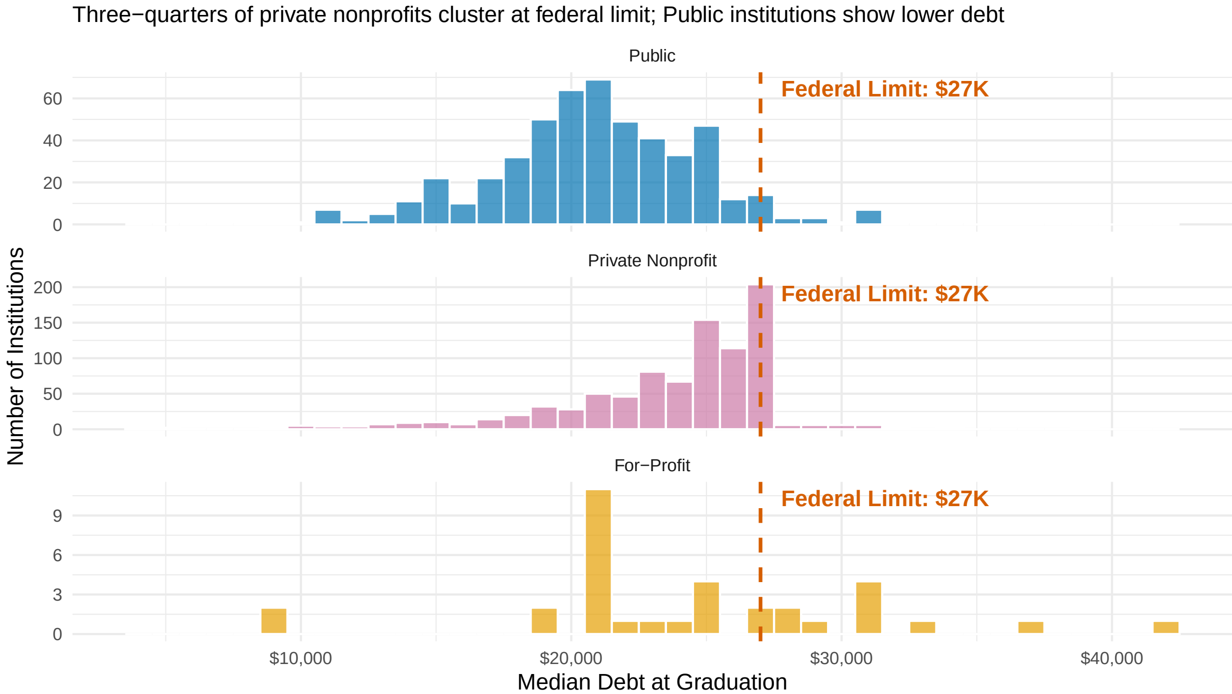
U.S. Department of Education, College Scorecard. (2025). *Most Recent Institution-Level Data.* Retrieved December 9, 2025, from https://collegescorecard.ed.gov/data/

APPENDIX

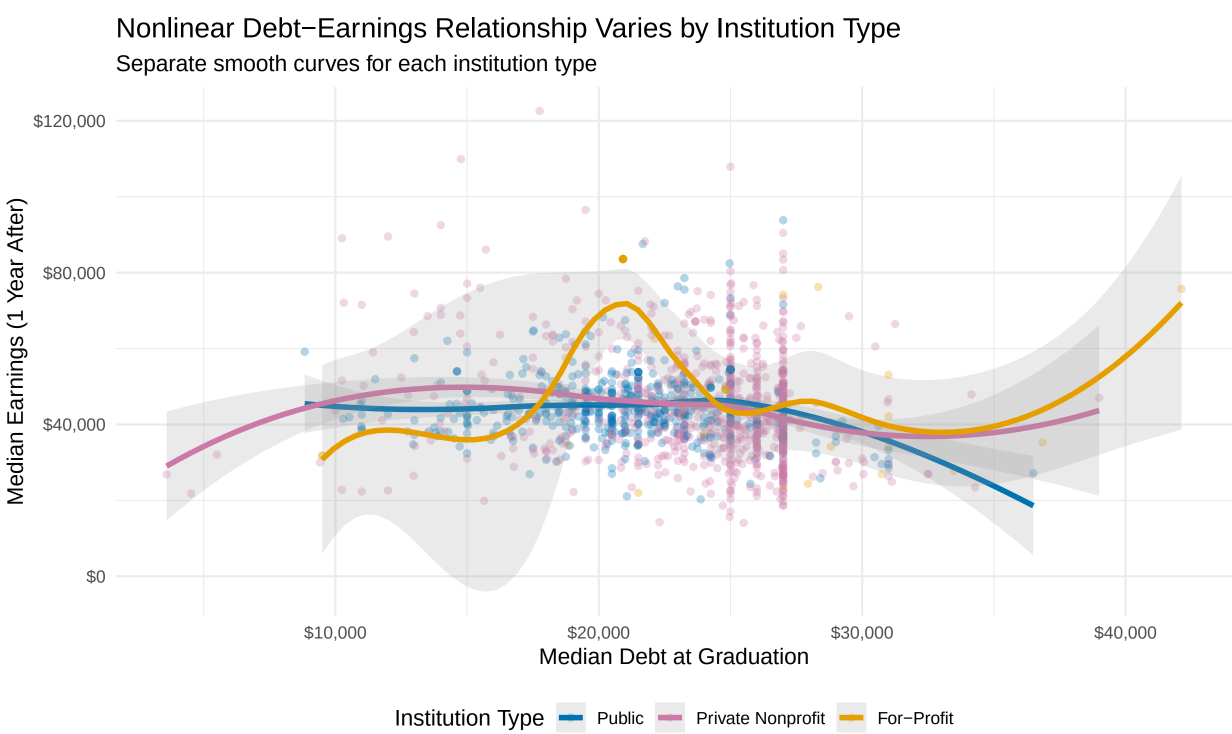
Zoomed View: Federal Loan Limit Clustering



Debt Distribution by Institution Type



Debt-Earnings Relationship by Institution Type



Why Graduation rate over Region?

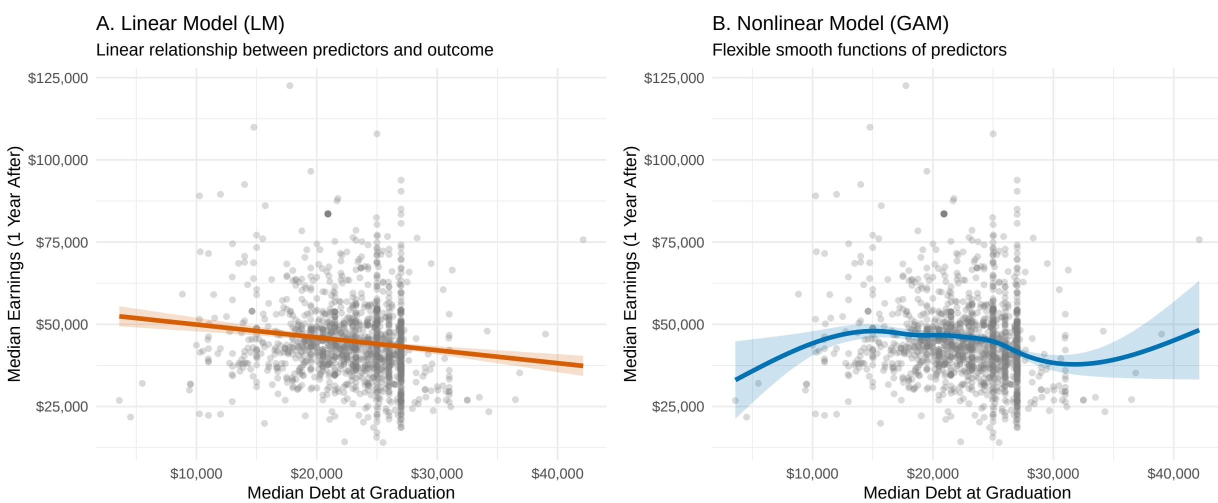
Table 14: Distribution by Region (Final Dataset)

Region	Count	Percentage (%)
Northeast	375	26.2
Midwest	386	27.0
South	473	33.1
West	196	13.7

Table 15: Why Focus on Graduation Rate: Variation in Debt-Earnings Correlation

Sector	Regional Range	Grad Rate Range	Ratio (Grad/Region)
Public	0.674	0.281	0.4×
Private Nonprofit	0.167	0.350	2.1×

Comparing Linear vs Nonlinear Approaches

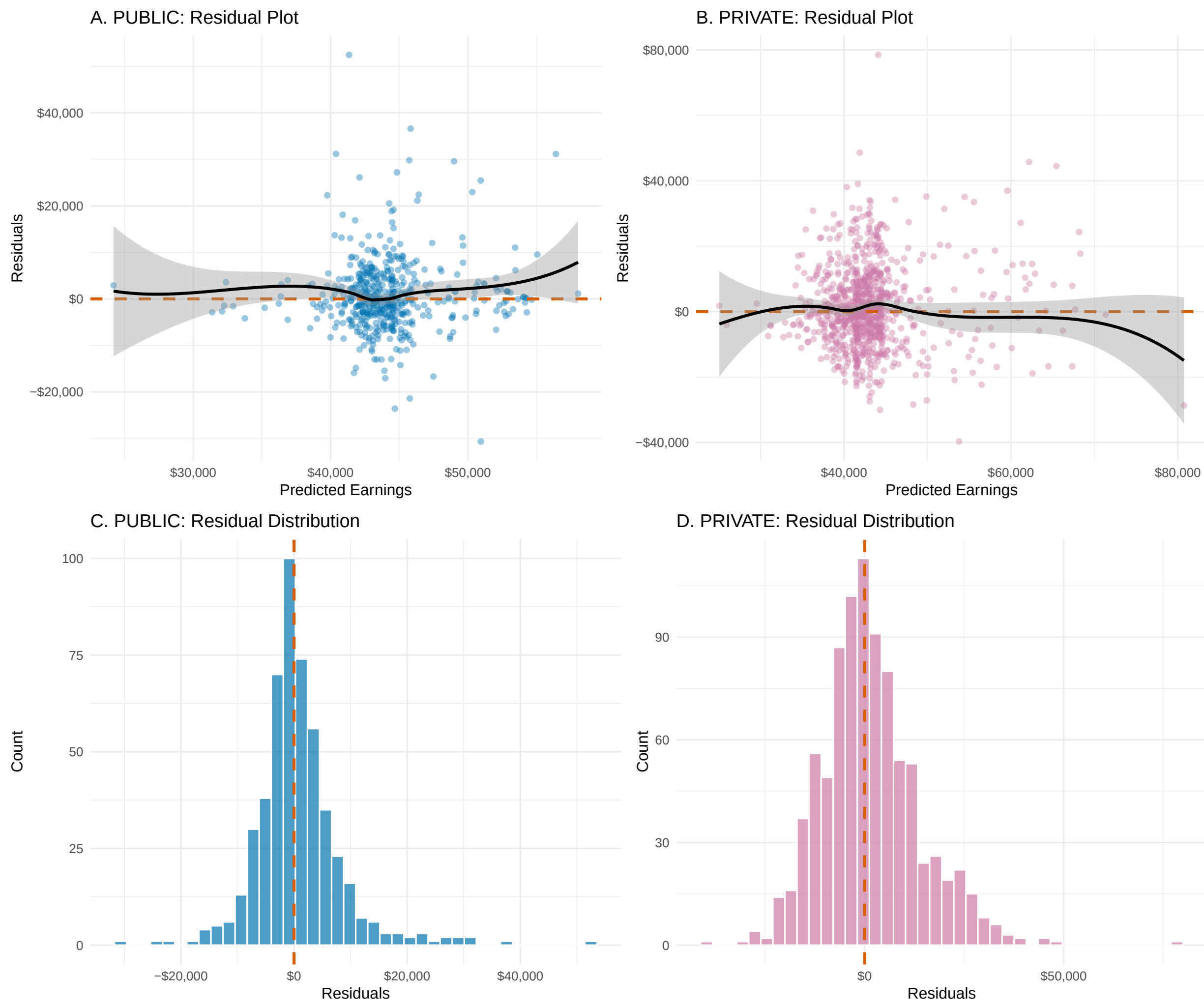


Model Complexity table (degrees of freedom)

Table 16: Model Complexity: Effective Degrees of Freedom (EDF)

Model	Total EDF	Sample Size	EDF per 100 obs
Public (Interaction)	17.5	507	3.45
Private (Both Smooth)	11.5	889	1.29

Residual Plots



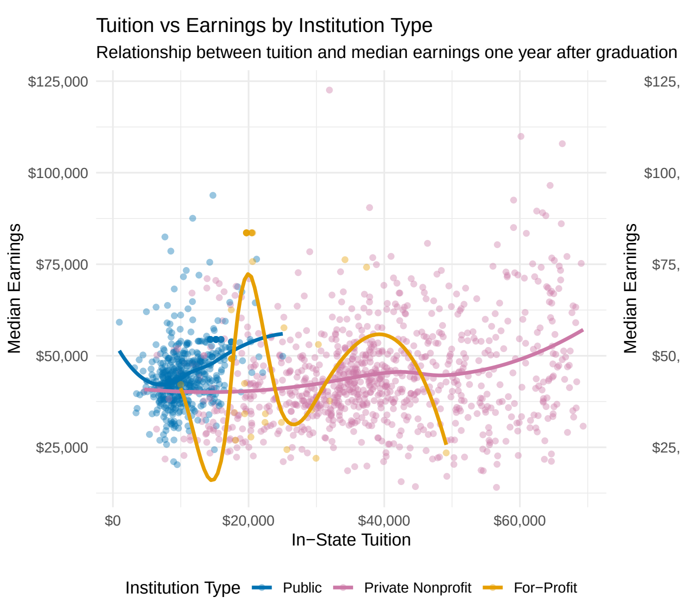
Full model summaries

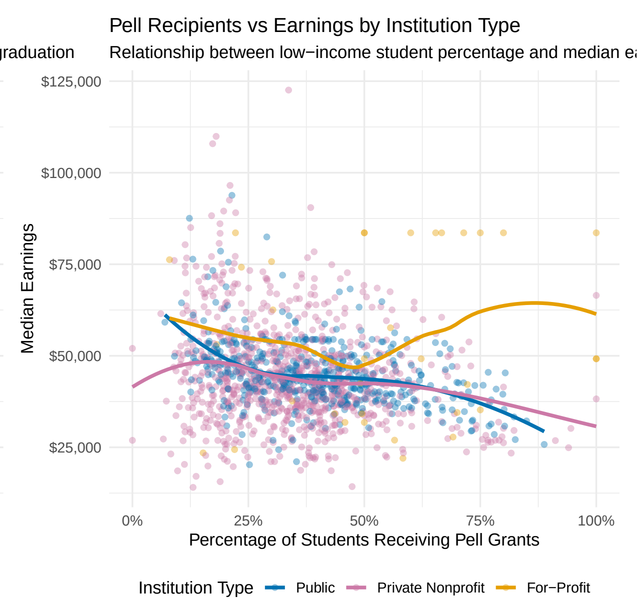
```
##
## Family: Scaled t(3.176,0.108)
## Link function: identity
##
## Formula:
## log_earnings ~ s(log_debt, grad_rate, k = 20)
##
## Parametric coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) 10.692219   0.005841    1830   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                          edf Ref.df Chi.sq p-value
## s(log_debt,grad_rate) 16.5  18.44  219.2  <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.194   Deviance explained = 22.2%
## -REML = -207.23  Scale est. = 1         n = 507
```

```
##
## Family: Scaled t(9.545,0.251)
## Link function: identity
##
## Formula:
## log_earnings ~ s(log_debt, k = 10) + s(grad_rate, k = 10)
##
## Parametric coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) 10.663595   0.009184    1161   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##               edf Ref.df Chi.sq  p-value
## s(log_debt)  4.001  5.008  23.64 0.000253 ***
## s(grad_rate) 6.470  7.603 117.61  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.142   Deviance explained = 14.2%
## -REML = 149.09  Scale est. = 1         n = 889
```

Unmeasured Confounding Variables





Outlier Visualization

Generative AI Use

AI assistance statement. We used OpenAI ChatGPT, Google Gemini, and Anthropic Claude for R code syntax generation and minor text edits; all analysis decisions and interpretations are our own.