Investigating UMass's Rising In-State Tuition

Madelaine Mildwoff, Ryan Nguyen, Aniket Patil, Mila Lewis-Peinado

May 2024

Abstract

This study examines the determinants of in-state undergraduate tuition at the University of Massachusetts Amherst from 1984 to 2021. Utilizing a multiple linear regression model, we analyze the impact of key variables including year, lagged enrollment, and log-transformed lagged state income tax revenue. Our findings indicate that tuition is significantly influenced by both university-specific factors, such as enrollment, and broader economic indicators, such as state income tax revenue. While the model captures important trends, limitations related to data availability and multicollinearity are acknowledged. Future research is recommended to include a broader range of variables and apply advanced econometric techniques.

1 Introduction

In the United States, as students graduate high school, many are encouraged to pursue higher education for the opportunity to increase their prospects in the job market. Despite this, undergraduate enrollment has not been on a steady incline. Undergraduate enrollment in the United States was growing rapidly in the late-1900s but then peaked in 2010 with approximately 21 million enrolled students. From 2010-2021, undergraduate enrollment slowly declined by about 2.3 million (Education Data Initiative, 2023). In 2021, only around 62% of high school graduates in the United States immediately enrolled in college for either a 2-year or 4-year program. Many people believe that the high costs of higher education play a role in deterring high school students from attending college. The annual average tuition rates for colleges in the United States have increased substantially since the 1900s. Even after adjusting for inflation, tuition rates have increased by about 747.8% in the 60 years between 1963 and 2023 (Education Data Initiative, 2023). This is a huge increase especially considering the fact that the US minimum wage when also adjusted for inflation has not had such a dramatic change. The decade with the highest increase was the 1980s when tuition went up by 52%. For the majority of these 60 years, tuition rates have been increasing annually but the rates have plateaued since 2019 (McGurran, 2023).

The University of Massachusetts at Amherst (UMass Amherst) is primarily funded by the state and has traditionally served in-state students. It isn't clear exactly how much of UMass Amherst's funding is from the state, but they've disclosed that ”a significant funding source is the Commonwealth of Massachusetts” (UMass Amherst, 2023). This past fall, 73% of UMass Amherst's incoming class are Massachusetts residents (Blaguszewski, 2023). UMass Amherst has grown a lot over the past 30 years – admission rates have gone down as they are receiving more applications than they can accept, which allows them to be more selective and competitive with who they admit. What sets UMass Amherst's growth in the 2000-2020s apart from their counterpart universities is their aggressive expansion of enrollment. From 2005 to 2014, UMass Amherst has increased enrollment by 27.3%, while New England state universities have seen a 1.7% increase and Massachusetts private 4-year institutions have seen a 11.8% increase (Sullivan et al., 2016). Out-of-state undergraduate enrollment has made up most of this increase (Sullivan et al., 2016), which benefits UMass Amherst because out-of-state tuition is a lot higher than in-state tuition. As a result, UMass Amherst is growing more dependent on out-of-state enrollment which led us to wonder about the changes in in-state tuition. Especially whether they are more affected by the increase in enrollment or changes in the Massachusetts' economy.

Like most universities across the nation (Hemelt 2013), UMass Amherst tuition has increased over the past decade for both in-state and out-of-state undergraduate students. The rates at which tuition for in-state and out-of-state undergraduate students has increased are generally close to each other, but not always identical. Most recently, for the 2023-2024 academic year, UMass Amherst increased the tuition for in-state undergraduates by 2.5% and for out-of-state undergraduates by 3%. But while UMass Amherst's enrollment is increasing, the number of high school graduates in Massachusetts is declining and graduation rates across the nation are stagnating (Murnane 2013); the Western Interstate Commission for Higher Education (WICHE) has predicted that in Massachusetts, "the number of high school graduates will decline by 11.4 percent from 2015-16 to 2027-28" (Sullivan et al., 2016). Massachusetts high school graduation rates are declining, yet general undergraduate enrollment and in-state tuition is still increasing. This led us to solidify our research question – is in-state tuition more affected by UMass Amherst's yearly enrollment or by statewide factors?

2 Methodology

This study employs data from two primary sources: the University of Massachusetts (UMass) Amherst Fact Sheet and the Federal Reserve Economic Data (FRED) provided by the Federal Reserve Bank of St. Louis. The analysis spans the period from 1984 to 2021, incorporating annual data related to UMass Amherst alongside corresponding statewide metrics for Massachusetts. The primary objective of this analysis is to investigate the determinants of in-state undergraduate tuition at UMass Amherst using a multiple linear regression model.

The dataset comprises 38 annual observations, each representing a year's worth of aggregated statistics for UMass Amherst and state-level characteristics for Massachusetts. The variables included in the analysis are: In-State Undergraduate Tuition in dollars ($totalis.ug$), Year ($year$), Bachelor's Degree Seeking Students Enrolled ($bat.enrolled$), and Total Income Tax Received by the State in dollars (mit).

To capture the delayed effects of enrollment and state income tax on tuition, the variables for bachelor's degree-seeking students enrolled and total income tax received by the state are included in the model with a one-year lag. This decision is based on the assumption that changes in these variables do not have an immediate impact on tuition rates but rather influence tuition in subsequent years. Thus, the lagged variables are: Lagged Bachelor's Degree Seeking Students Enrolled ($bat.enrolled_{t_{-1}}$), representing the number of bachelor's degree-seeking students enrolled in the previous year, and Log-Transformed Lagged Total Income Tax ($\log(mit_{t_{-1}})$), representing the natural logarithm of the total income tax received by the state in the previous year.

To justify the transformation of the *mit* variable (Total Income Tax Received by the State), we conducted an analysis comparing the distribution of the original *mit* values to their logarithmic transformation $\log($*mit*$)$. This analysis included visual inspections of histograms and Q-Q plots, as well as formal statistical testing using the Shapiro-Wilk test for normality.

Figure 1 presents the histograms and Q-Q plots for both the original *mit* and the transformed $\log($*mit*$)$ variables. The histogram of the original *mit* values indicates a right-skewed distribution, with a significant concentration of values on the lower end of the range. In contrast, the histogram of $\log($*mit*$)$ shows a more symmetrical distribution, suggesting an improvement towards normality.

The Q-Q plot for the original *mit* variable further supports the presence of skewness, as the points deviate significantly from the reference line, especially at the tails. On the other hand, the Q-Q plot for $\log($*mit*$)$ demonstrates that the points are more closely aligned with the reference line, indicating that the log-transformed data more closely approximates a normal distribution.

To quantitatively assess the normality of the distributions, we performed the Shapiro-Wilk test. For the original *mit* variable, the Shapiro-Wilk test yielded a W statistic of 0.9524 and a p-value of 0.1154. Although this p-value is greater than the conventional significance level of 0.05, indicating that we cannot reject the null hypothesis of normality, the W statistic suggests some deviation from a perfect normal distribution. In contrast, the Shapiro-Wilk test for the $\log($*mit*$)$ variable produced a W statistic of 0.9706 and a p-value of 0.4261, providing stronger evidence that the log-transformed data does not deviate significantly from normality.

Based on these findings, the logarithmic transformation of the *mit* variable is justified as it results in a distribution that better approximates normality. This transformation helps satisfy the assumptions of linear regression, particularly the assumption of normally distributed residuals, which enhances the reliability and validity of the regression model's estimates and inferences.

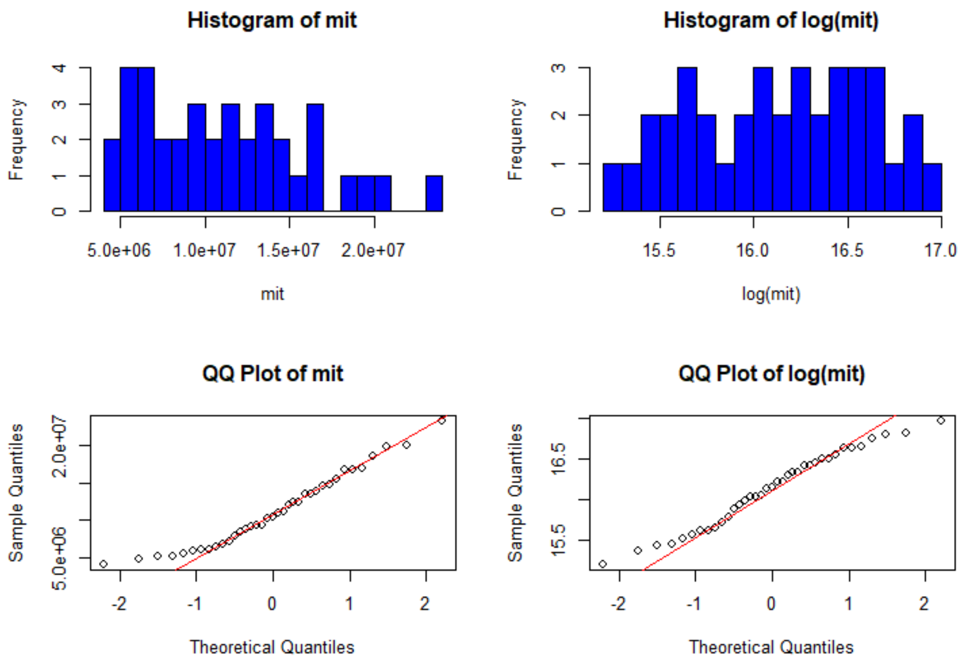

Figure 1: Effect of Log Transformation on the Distribution of *mit*

The multiple linear regression model is formulated as follows:

$$totalis.ug = \beta_0 + \beta_1 \cdot year + \beta_2 \cdot bat.enrolled_{t_{-1}} + \beta_3 \cdot \log(mit_{t_{-1}}) + \varepsilon$$

where $totalis.ug$ represents the in-state undergraduate tuition, $year$ is the year of observation, $bat.enrolled_{t_{-1}}$ is the lagged number of bachelor's degree-seeking students enrolled, $\log(mit_{t_{-1}})$ is the log-transformed lagged total income tax received by the state, β_0 is the intercept, β_1, β_2, and β_3 are the coefficients for the respective predictors, and ε is the error term.

The primary objective is to investigate the relationship between annual tuition rates and a series of explanatory variables that encompass both university-specific metrics and broader economic indicators for Massachusetts. The inclusion of lagged variables is based on the assumption that their effects on undergraduate tuition are not contemporaneous, thus requiring a lag to capture these delayed impacts.

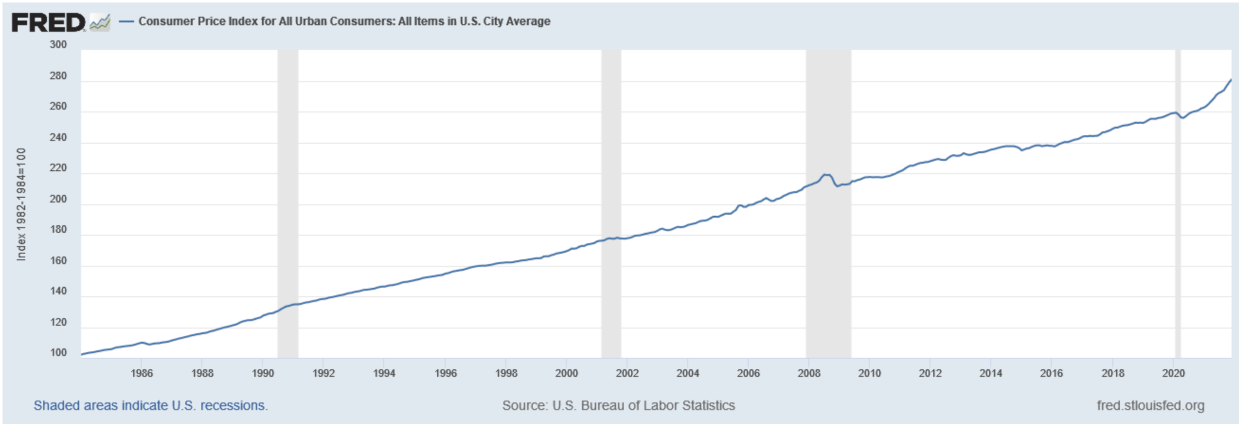

Figure 2: Consumer Price Index for All Urban Consumers: All Items in U.S. City Average (1984-2021). Source: U.S. Bureau of Labor Statistics, retrieved from FRED, Federal Reserve Bank of St. Louis; `https://fred.stlouisfed.org/series/CPIAUCSL`, May 17, 2024.

Figure 2 illustrates the Consumer Price Index (CPI) for All Urban Consumers, covering all items in U.S. city averages from 1984 to 2021. The CPI shows a steady increase over the years, indicating a persistent rise in the inflation rate during this period. This trend is relevant to our study as the year variable in our regression model captures this temporal increase in inflation, which can influence the in-state undergraduate tuition rates at UMass Amherst.

Our analysis aims to determine whether in-state undergraduate tuition rates at UMass Amherst are predominantly influenced by enrollment metrics or if broader state-level characteristics also play a significant role. Specifically, we examine the determinants of tuition adjustments over time, considering the interplay between university enrollment figures and the state's economic and demographic landscape.

The predictors selected for the regression include variables directly related to enrollment, such as the number of bachelor's degrees awarded and the number of bachelor's degree-seeking students enrolled. State characteristics are included to capture demographic and economic factors. Room and board costs were excluded from the analysis due to their susceptibility to fluctuations in local cost of living. Additionally, the scope of our investigation is intentionally focused on in-state undergraduate tuition to maintain relevance to the segment of the student population directly impacted by state residency status. Consequently, variables related to out-of-state and graduate tuition rates, as well as associate and graduate degree enrollment and awards, are omitted. These populations are governed by distinct pricing and funding mechanisms that are not directly pertinent to the research question at hand.

3 Results

3.1 Descriptive Statistics

This section presents the findings from the regression analysis conducted to investigate the determinants of in-state undergraduate tuition at the University of Massachusetts Amherst. The results are organized by the impact of the key variables: year, bachelor's degree enrollment and total income tax received.

Figure 3 displays the histogram of in-state undergraduate tuition (totalis.ug) over the period from 1985 to 2021. The distribution of the outcome variable is right-skewed, indicating that while a significant number of observations cluster around a lower tuition range, there are a few instances where the tuition values are substantially higher.

The right-skewed distribution suggests that the majority of tuition rates are concentrated around lower values, with fewer instances of higher tuition rates. This skewness could be attributed to periods of stability or only marginal increases in tuition rates over time. When tuition rates remain stable or increase gradually, they tend to accumulate around certain values, creating a peak.

Additionally, the histogram shows notable peaks and gaps, indicating that tuition rates may have experienced distinct phases of rapid increases followed by periods of stability. This pattern reflects the historical changes in tuition rates, where certain years may have seen substantial increases due to economic or policy changes, while other years remained relatively stable

Table 1 presents the summary statistics for the key variables included in the regression analysis: in-state undergraduate tuition (*totalis.ug*), year, lagged bachelor's degree-seeking students enrolled (*bat.enrolled_lag1*), and the log-transformed lagged total income tax received by the state

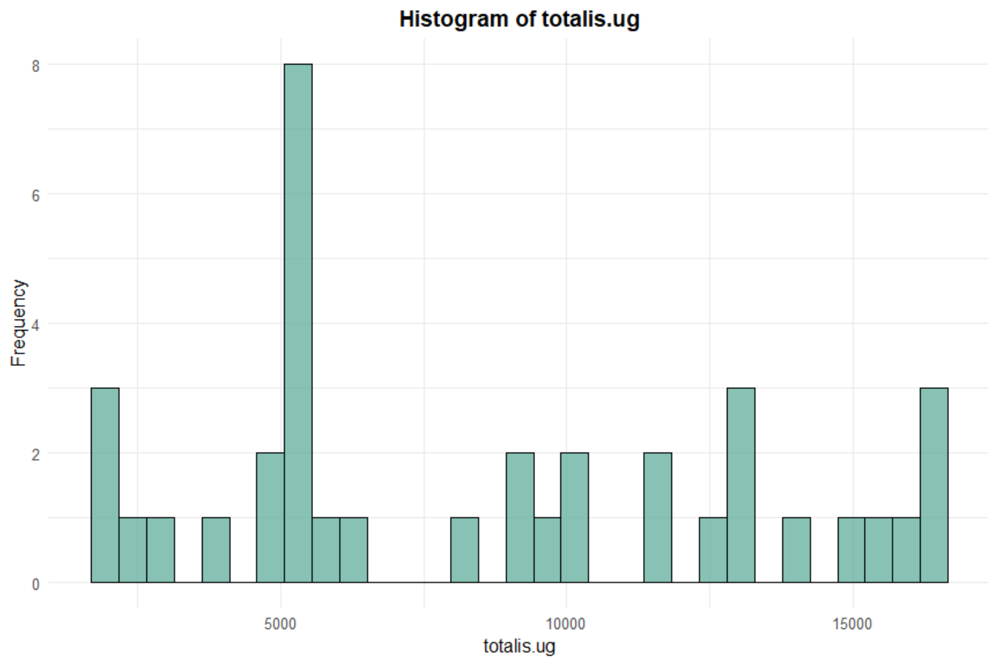

Figure 3: Histogram of In-State Undergraduate Tuition (totalis.ug)

(*log_mit_lag1*). The in-state undergraduate tuition ranges from $1,947 to $16,439, with a mean of $8,706 and a median of $8,232. The first quartile is $5,212, and the third quartile is $13,230, indicating a relatively wide spread in tuition values. The dataset spans from 1985 to 2021(as we lagged the data by 1 year), with the median year being 2003.

The number of lagged bachelor's degree-seeking students enrolled ranges from 16,885 to 24,167, with a mean of 19,770 and a median of 19,446. The first quartile is 17,949, and the third quartile is 21,645, showing a significant range in student enrollment numbers. The log-transformed lagged total income tax received by the state ranges from 15.07 to 16.82, with a mean of 16.08 and a median of 16.13. The first quartile is 15.66, and the third quartile is 16.45, indicating moderate variability in state income tax revenue.

Figure 4 shows the scatterplot matrix of the variables included in the regression analysis. Each plot in the matrix helps visualize the relationships between pairs of variables, and the correlation coefficients are also displayed.

The scatterplot matrix reveals several key relationships. There is a strong positive correlation ($r = 0.982$) between in-state undergraduate tuition (*totalis.ug*) and year, indicating that tuition

	totalis.ug	year	bat.enrolled_lag1	log_mit_lag1
Min.	1947	1985	16885	15.07
1st Qu.	5212	1994	17949	15.66
Median	8232	2003	19446	16.13
Mean	8706	2003	19770	16.08
3rd Qu.	13230	2012	21645	16.45
Max.	16439	2021	24167	16.82

Table 1: Summary Statistics for totalis.ug, year, bat.enrolled_lag1, and log_mit_lag1

has increased consistently over time. This trend is visually represented by the upward trend in the scatterplots. Additionally, a positive correlation ($r = 0.830$) is observed between tuition and the number of lagged bachelor's degree-seeking students enrolled (*bat.enrolled_lag1*), suggesting that higher enrollment numbers are associated with higher tuition rates. The positive correlation ($r = 0.783$) between year and lagged enrollment numbers shows that student enrollment has generally increased over the years, reflecting the university's growth. Lastly, the positive correlation ($r = 0.982$) between year and log-transformed state income tax revenue suggests that state income tax revenue has also increased over time.

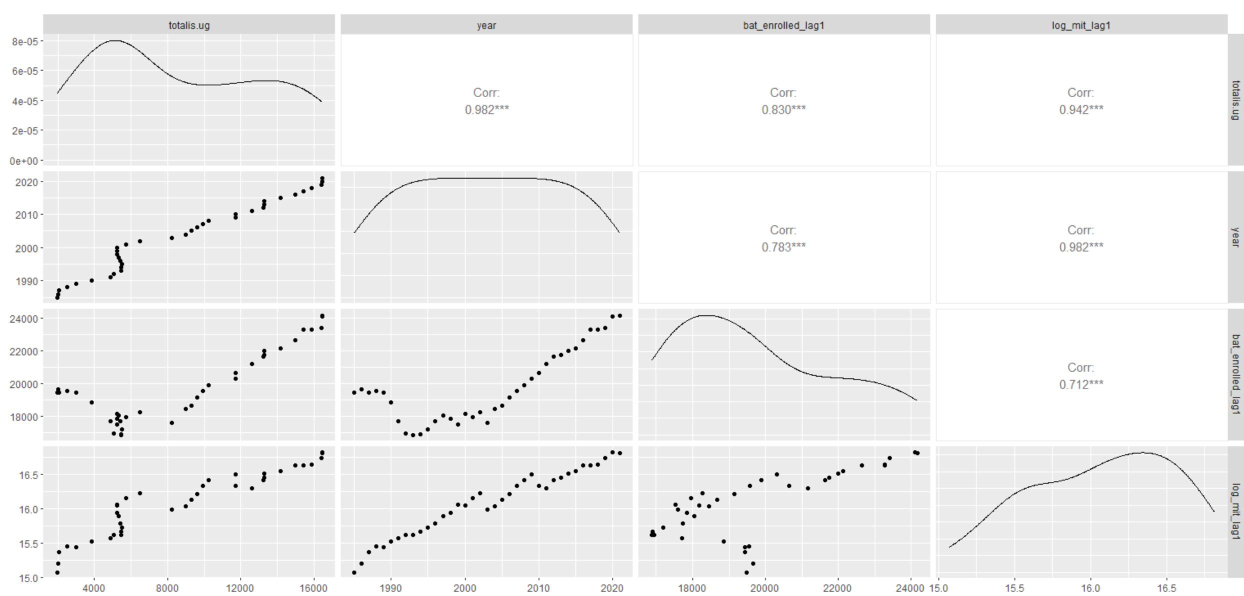

Figure 4: Scatterplot Matrix of totalis.ug, year, bat.enrolled_lag1, and log_mit_lag1

3.2 Robustness Checks

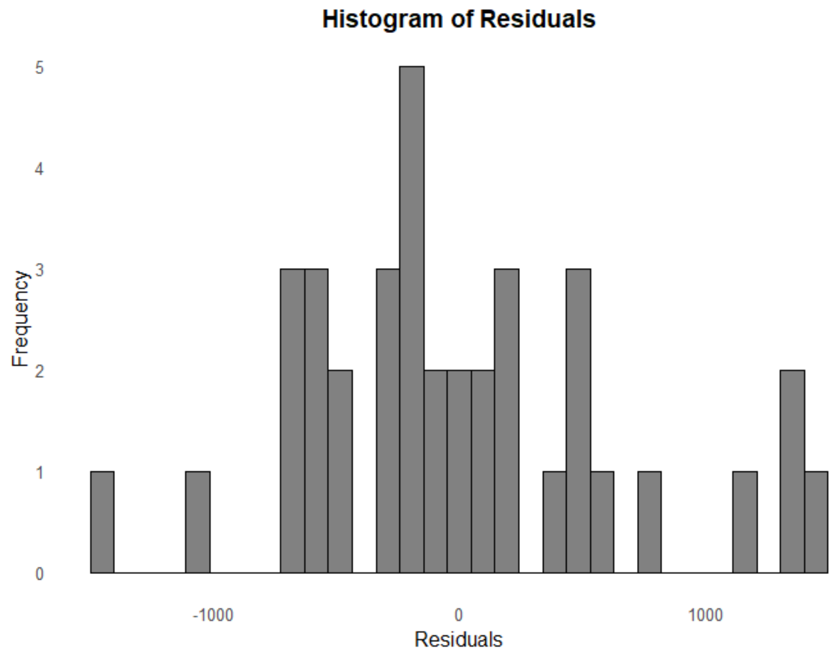

Figure 5: Histogram of Residuals

Figure 5 presents the histogram of the residuals from the regression model. The residuals are approximately centered around zero, indicating that the model does not systematically overestimate or underestimate the in-state undergraduate tuition. The distribution of the residuals appears roughly symmetric but shows some deviations from normality, with several peaks and gaps. This suggests that while the model captures the general trend in the data, there may be some outliers or variations not fully accounted for by the explanatory variables.

Figure 6 presents the boxplot of the residuals from the regression model. The boxplot shows that the residuals are symmetrically distributed around zero, with the median close to zero, indicating that the model does not have a significant bias in its predictions. The interquartile range (IQR) is relatively narrow, suggesting that most of the residuals are close to the median. However, there are a few extreme values that fall outside the whiskers, indicating the presence of some outliers.

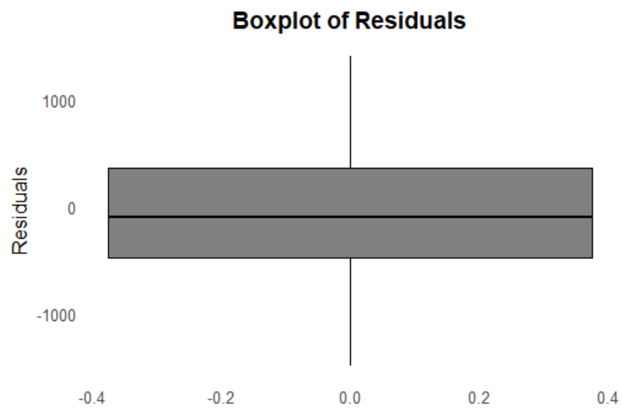

Figure 6: Boxplot of Residuals

Figure 7 presents the Q-Q plot of the residuals from the regression model. The Q-Q plot compares the sample quantiles of the residuals to the theoretical quantiles of a standard normal distribution. Most of the residuals fall along the reference line, indicating that they are approximately normally distributed. However, there are some deviations at the tails, suggesting the presence of a few outliers and slight departures from normality.

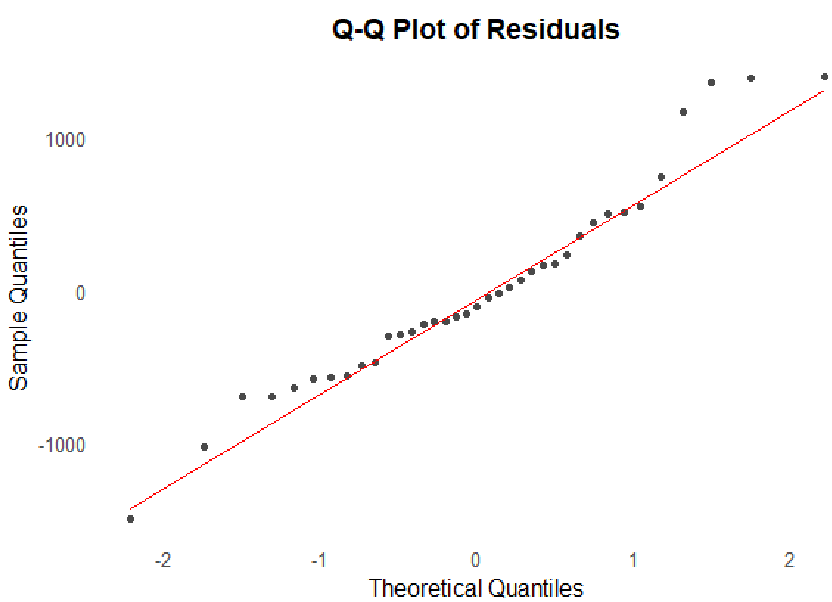

Figure 7: Q-Q Plot of Residuals

The residual plot indicates that our first assumption is met (the linear model), so our estimators are consistent. The sample size is relatively large, so our estimators should be close to the estimand. The boxplot shows the median close to zero, indicating that the model does not have a bias in underestimating or overestimating the dependent variable. The residuals are symmetric and approximately normally distributed, confirming that our estimators are t-distributed.

Figure 8 presents the residuals vs. fitted values plot from the regression model. The plot displays the residuals on the y-axis and the fitted values on the x-axis. The residuals are scattered randomly around the horizontal line at zero, which is indicated by the red dashed line. This random scatter suggests that the assumption of linearity is reasonable, and there is no apparent pattern indicating a systematic relationship between the residuals and the fitted values.

The mean of the residuals is approximately zero, as evidenced by their symmetric distribution around the horizontal line at zero. This indicates that the model does not systematically overestimate or underestimate the dependent variable. While the variance does change, it is constant from 0 to 7000 and from 7500 onwards. This constancy in variance within these ranges suggests that the variability in the residuals is relatively stable, further supporting the validity of the linear regression model and its assumptions.

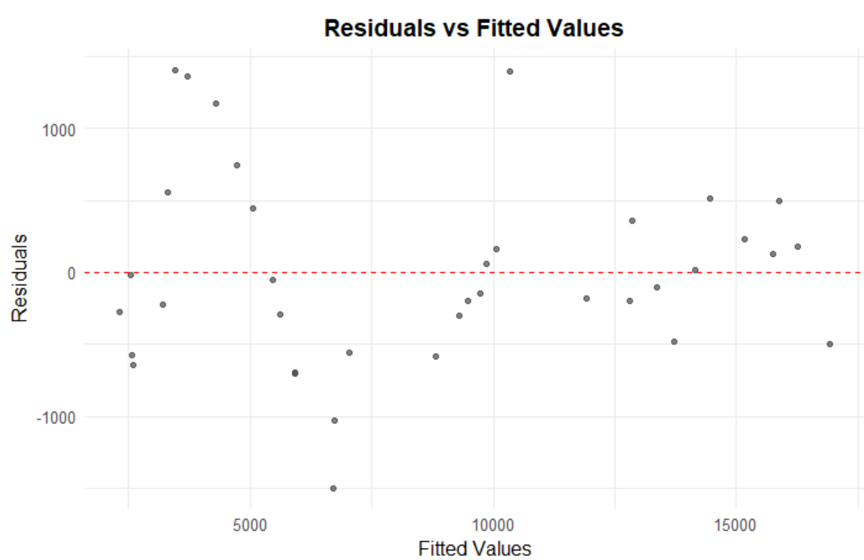

Figure 8: Residuals vs. Fitted Values

3.3 Regression Analysis

Table 2 presents the regression results for the model predicting *totalis.ug* based on the year, lagged bat enrollment (*bat_enrolled_lag1*), and lagged logarithm of MIT values (*log_mit_lag1*). The intercept is estimated at -1,168,000 with a standard error of 123,100. The t-value for the intercept is -9.481, and the corresponding p-value is less than 0.001, indicating that the intercept is significantly different from zero at the 1% significance level. This large negative intercept suggests that, at the baseline (when all predictors are zero), *totalis.ug* would be significantly low, though this baseline condition may not be practically meaningful due to the nature of the predictors.

The coefficient for *year* is estimated to be 627.4 with a standard error of 73.60. The t-value is 8.524, and the p-value is less than 0.001, indicating that the coefficient is highly significant at the 1% level. This result suggests that, holding other variables constant, each additional year is associated with an increase in *totalis.ug* by approximately $627.4 . This positive trend implies that *totalis.ug* has been increasing over time.

The coefficient for *bat_enrolled_lag1* is estimated at 0.1741 with a standard error of 0.0963. The t-value for this coefficient is 1.808, and the p-value is 0.0798, indicating marginal significance at the 10% level. This suggests that, holding other variables constant, an increase of 1 unit in lagged bat enrollment is associated with an increase in *totalis.ug* by approximately $0.1741. This positive relationship indicates that higher bat enrollment in the previous period tends to be associated with slightly higher *totalis.ug*.

The coefficient for *log_mit_lag1* is estimated at -5,221 with a standard error of 1,480. The t-value is -3.527, and the p-value is 0.00126, indicating that this coefficient is significant at the 1% level. Interpreting the coefficient of a log-transformed variable, this result suggests that, holding other variables constant, a 1% increase in the lagged MIT values is associated with a decrease in *totalis.ug* by approximately $52.21. This negative relationship implies that higher lagged MIT values are associated with lower *totalis.ug*, which could reflect a compensatory or balancing mechanism in the underlying dynamics of the system.

Variable	**Estimate**	**Std. Error**	**t value**	**Pr($>|t|$)**
(Intercept)	-1,168,000	123,100	-9.481	<0.001***
year	627.4	73.60	8.524	<0.001***
bat_enrolled_lag1	0.1741	0.0963	1.808	0.0798.
log_mit_lag1	-5,221	1,480	-3.527	0.00126**

Table 2: Regression Results

3.4 Model Fit Statistics

Table 3 presents the key statistics for assessing the fit of the regression model. The residual standard error is 687.3, based on 33 degrees of freedom, indicating the average distance that the observed values fall from the regression line.

The multiple R-squared value of 0.9807 suggests that approximately 98.07% of the variance in the variable, in-state undergraduate tuition, is explained by the model. This high R-squared value indicates a strong relationship between the predictors and the dependent variable, meaning the model provides a good fit to the data.

The adjusted R-squared value, which adjusts for the number of predictors in the model, is 0.9789. This value is slightly lower than the multiple R-squared, reflecting the adjustment for the complexity of the model. Nevertheless, it still indicates that a large proportion of the variance in tuition is explained by the model, confirming the robustness of the model fit even when accounting for the number of predictors.

The F-statistic for the model is 558.1, with 3 and 33 degrees of freedom. This statistic tests the null hypothesis that all regression coefficients are equal to zero. The associated p-value is less than 2.2e-16, and indicates that the model is highly significant. This provides strong evidence against the null hypothesis, suggesting that the predictos are significantly related to in-state undergraduate tuition.

Metric	Value
Residual standard error	687.3 on 33 degrees of freedom
Multiple R-squared	0.9807
Adjusted R-squared	0.9789
F-statistic	558.1 on 3 and 33 DF
p-value	$< 2.2e-16$

Table 3: Model Fit Statistics

4 Discussion

4.1 Interpretation

From our research, we concluded that the cost of tuition at UMass Amherst is influenced by a combination of both yearly enrollment and broader statewide economic factors. The regression analysis identified year, enrollment, and total state income tax as significant predictors of tuition, with coefficients statistically different from zero. Initially, median household income was included in the regression model, but despite showing a correlation with tuition, it was excluded due to its high p-value, indicating a lack of statistical significance.

The coefficient for the year variable is 563.94, indicating that for each additional year, holding all other variables constant, the tuition increases by approximately $563.94. This significant positive trend reflects the broader inflationary pressures and increasing costs associated with higher education over the years. The steady rise in tuition over time can be attributed to factors such as inflation, increased operational costs, and the need for the university to expand and improve its facilities and services.

The enrollment variable, represented as the lagged number of bachelor's degree-seeking students, has a coefficient of 0.17. Although this effect may appear small on a per-student basis, it becomes substantial when considering large changes in enrollment numbers. The positive relationship between enrollment and tuition could be driven by the increased demand for university resources and the need to maintain or enhance the quality of education and services provided to a growing student body.

The total state income tax, log-transformed and lagged by one year, has a coefficient of -

5060.75. This negative relationship suggests that higher state income tax revenue allows for greater state funding or subsidies to the university, thereby reducing the financial burden on students in the form of lower tuition rates. The significance of this variable underscores the importance of state economic health and policy decisions in influencing higher education costs.

The distribution of the residuals from our regression model was approximately normal with a mean close to zero, indicating that the model does not systematically overestimate or underestimate tuition. However, the analysis revealed that the variance of the residuals was not consistent throughout the years, a phenomenon known as heteroscedasticity. This variability in residuals suggests that the model's accuracy in predicting tuition costs may fluctuate over time, potentially reducing the reliability of future projections. Inconsistent variance in residuals could be due to various factors, such as changing economic conditions, policy shifts, or unexpected fluctuations in university expenses and funding.

The findings from this analysis provide valuable insights into the factors driving tuition costs at UMass Amherst. The significant impact of the year variable highlights the inevitable rise in tuition over time, influenced by long-term economic trends. The positive correlation with enrollment emphasizes the resource demands placed on the university by increasing student numbers. Meanwhile, the negative relationship with state income tax revenue underscores the critical role of state funding in mitigating tuition hikes.

While the model successfully identifies key determinants of tuition, the presence of heteroscedasticity indicates that additional factors or more sophisticated modeling techniques might be necessary to improve the precision of future tuition predictions. Policymakers and university administrators should consider these findings when planning budget allocations and setting tuition rates, ensuring that they account for both immediate enrollment pressures and broader economic conditions.

4.2 Limitations

Our study faced several limitations, primarily related to the availability and quality of data. The dataset used for the analysis consisted of only 38 annual observations, spanning from 1984 to 2021. This limited sample size restricts the generalizability of our findings and may reduce the statistical power of our regression analysis. A larger dataset with more observations could potentially provide a more robust analysis and yield more reliable results.

One significant challenge was the high correlation between the variables included in the regression model. Multicollinearity can inflate the standard errors of the coefficient estimates, making it difficult to determine the individual effect of each predictor on the dependent variable. In our analysis, the high correlation between variables such as year, enrollment, and state income tax likely introduced redundancy in the regression model. This multicollinearity can lead to less stable coefficient estimates and can obscure the true relationships between the predictors and the dependent variable.

The selection of variables was also constrained by the availability of publicly accessible data. While we included key variables such as in-state undergraduate tuition, year, enrollment, and state income tax, other potentially important factors were excluded due to data limitations. For instance, we initially considered including median household income as a predictor. Although it showed a correlation with tuition, it was ultimately excluded due to its high p-value, indicating a lack of statistical significance. This exclusion may have omitted an important aspect of the economic environment that could influence tuition rates.

Additionally, the data for certain variables were lagged to capture delayed effects on tuition. While this approach helps address issues of contemporaneous impact, it also introduces complexity and potential inaccuracies in the timing of these effects. The assumption that the impact of variables such as enrollment and state income tax is fully realized after one year may not hold true in all cases, leading to potential misestimations of their influence.

Another limitation is related to the heteroscedasticity observed in the residuals of our regression model. The inconsistent variance of residuals across different years suggests that the model's accuracy in predicting tuition costs may vary over time. This heteroscedasticity violates one of the key assumptions of linear regression and could affect the reliability of our coefficient estimates and hypothesis tests. Future research could benefit from employing more advanced econometric techniques, such as generalized least squares (GLS) or robust standard errors, to address this issue.

Finally, the external validity of our findings is limited to the context of UMass Amherst and may not be directly applicable to other universities or states with different economic conditions and educational policies. The unique characteristics of UMass Amherst, such as its funding structure and enrollment patterns, may influence the generalizability of our results.

While our study provides valuable insights into the determinants of in-state undergraduate tuition at UMass Amherst, the limitations related to data availability, multicollinearity, variable selection, and heteroscedasticity highlight the need for cautious interpretation of our findings. Future research with larger datasets, additional variables, and advanced econometric techniques could help address these limitations and provide a more comprehensive understanding of the factors influencing tuition rates.

4.3 Future Research

Future research could explore several additional topics to build on our findings. One area of interest is the comparative analysis of in-state versus out-of-state tuition, examining the factors that contribute to differences in these tuition rates. Another important direction is to employ alternative econometric models that can better handle the issue of multicollinearity among highly correlated covariates. Additionally, it would be valuable to investigate how the tuition trends at UMass Amherst compare to those at other institutions, including both private and state universities. Finally, incorporating variables that were not available in our current study, such as construction costs, financial aid availability, and the median household income of students, could provide a more comprehensive understanding of the determinants of tuition rates.

5 Conclusion

This study aimed to investigate the determinants of in-state undergraduate tuition at the University of Massachusetts Amherst, focusing on the impact of yearly enrollment figures and broader statewide economic factors. Utilizing a multiple linear regression model, we analyzed 38 annual observations from 1984 to 2021, incorporating key variables such as in-state undergraduate tuition, year, lagged enrollment numbers, and log-transformed lagged state income tax revenue.

Our findings indicate that the cost of tuition at UMass Amherst is influenced by both university-specific and state-level factors. Specifically, the year variable demonstrated a significant positive trend, reflecting the broader inflationary pressures and rising costs associated with higher education over time. The lagged enrollment numbers also showed a positive relationship with tuition, suggesting that increasing student numbers contribute to higher tuition rates due to the demand for university resources and services. Additionally, the negative relationship between log-transformed lagged state income tax revenue and tuition underscores the role of state funding in mitigating tuition increases, highlighting the importance of state economic health and policy decisions.

Despite the robustness of our regression model, several limitations were identified. The small sample size and high correlation between variables posed challenges, potentially affecting the stability and reliability of our estimates. The exclusion of important variables such as median household income, construction costs, and financial aid availability due to data constraints also limits the comprehensiveness of our analysis. Furthermore, the presence of heteroscedasticity in the residuals suggests that the model's accuracy in predicting tuition costs may vary over time, indicating the need for more advanced econometric techniques in future research.

Looking ahead, future research could address these limitations by incorporating larger datasets, employing alternative models to manage multicollinearity, and including additional variables that capture a wider range of factors influencing tuition rates. Comparative analyses between in-state and out-of-state tuition, as well as studies examining tuition trends at other institutions, could also provide valuable insights.

Our study provides a detailed examination of the factors driving in-state undergraduate tuition at UMass Amherst, offering important implications for policymakers and university administrators. By understanding the interplay between enrollment metrics and state-level economic conditions, stakeholders can make more informed decisions regarding budget allocations and tuition rate adjustments, ultimately ensuring the sustainability and accessibility of higher education for future generations.

References

[1] Sullivan, G. W., Blackbourn, M., and Corvese, L. (2016). UMass at a Crossroads. Part 1: Is the UMass Enrollment Expansion Plan Sustainable?: White Paper No. 145. Distributed by ERIC Clearinghouse. https://files.eric.ed.gov/fulltext/ED598651.pdf

[2] Murnane, R. J. (2013). U.S. High School Graduation Rates: Patterns and Explanations. Journal of Economic Literature, 51(2), 370–422. http://www.jstor.org/stable/23644749

[3] Hemelt, S.W., and Marcotte, D.E. (2016). The Changing Landscape of Tuition and Enrollment in American Public Higher Education. RSF: The Russell Sage Foundation Journal of the Social Sciences 2(1), 42-68. https://doi.org/10.1353/rus.2016.0002.

[4] Sullivan, G., Blackbourn, M., and Corvese, L. (2016). UMass at a Crossroads. Part 3: UMass' Growing Dependency on Tuition and Fees and Strategic Recruitment of Out-of-State Students. White Paper No. 147. Distributed by ERIC Clearinghouse. https://files.eric.ed.gov/fulltext/ED598661.pdf

[5] Hanson, Melanie. "College Tuition Inflation Rate" EducationData.org, August 13, 2023, https://educationdata.org/college-tuition-inflation-rate

[6] McGurran, Brianna. "College Tuition Inflation: Compare the Cost of College over Time." Forbes, Forbes Magazine, 6 Nov. 2023, www.forbes.com/advisor/student-loans/college-tuition-inflation/.

[7] National Center for Education Statistics. (2023). Immediate College Enrollment Rate. Condition of Education. U.S. Department of Education, Institute of Education Sciences. Retrieved March 24, 2023, from https://nces.ed.gov/programs/coe/indicator/cpa.

[8] Blaguszewski, E. (2023). UMass Amherst welcomes 5,260 first-year students to campus. UMass Amherst. https://www.umass.edu/news/article/umass-amherst-welcomes-5260-first-year-students-campus

[9] Public disclosure/student consumer information. UMass Amherst. (2023). https://www.umass.edu/gateway/disclosure