When is it enough?

Tax Revenue Prediction

Christopher Guzman Silva

City College of New York

ABSTRACT

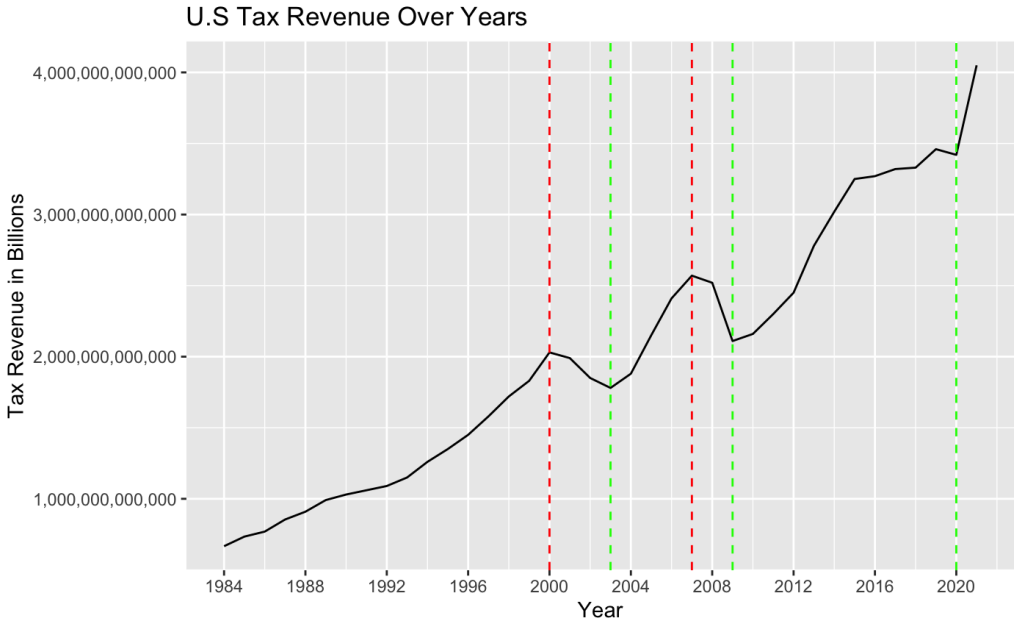
This project attempts to develop a predictive model for annual tax revenue in the US, that can be applied to other regions, by leveraging historical data and relevant economic and demographic indicators. The objective is to provide policymakers and stakeholders with a tool to anticipate tax revenue trends based on key factors influencing the fiscal landscape. The dataset comprises variables such as Gross Domestic Product (GDP), population, unemployment rate, median household income, and the Consumer Price Index (CPI). These indicators are hypothesized to play a significant role in shaping tax revenue dynamics. The project methodology involves data cleaning, exploratory data analysis, and the construction of a multi-linear regression model using the R programming language. The model aims to capture the relationships between the selected independent variables and the dependent variable, tax revenue. Through rigorous evaluation, the model's performance will be assessed, and insights will be drawn regarding the significance and impact of each variable. Interpretation of the model coefficients will shed light on the relative influence of economic and demographic factors on tax revenue fluctuations. Visualizations, including scatter plots and trend lines, will aid in conveying the findings to a diverse audience. The project's conclusion will summarize key insights, outline the factors most crucial in predicting tax revenue, and provide actionable recommendations for policymakers to enhance fiscal planning and decision-making. This Tax Revenue Prediction project seeks to contribute valuable insights into the intricate relationship between economic and demographic indicators and tax revenue.

I. Overview

This overview presents a historical analysis of tax revenue in the United States spanning the years 1984 to 2021. The dataset includes annual tax revenue figures, providing insights into the financial landscape of the country from this time. To better understand these trends, a graph 1[1] below was created to visualize the year-wise fluctuations and highlight significant points in the dataset.

Tax revenue showcases a dynamic interplay of factors, resulting in pronounced year-to-year fluctuations. This inherent responsiveness reflects the intricate relationship between economic shifts and tax policy adjustments. These fluctuations highlight the adaptability and complexity of the U.S. tax system, emphasizing the need for a flexible and responsive fiscal approach to navigate changing circumstances.

Graph 1: U.S Tax Revenue Over Years 1984-2021



The dataset unveils a compelling upward trajectory in tax revenue from 2010 to 2019 as you can see by the added vertical green lines, reaching its peak at $4.05 trillion in 2021. This decade-

[1] "U.S Tax Revenue Over Years 1963-2021" Amadeo, K. (2022, December 12). *U.S. Federal Government Tax Revenue*. The Balance. https://www.thebalancemoney.com/current-u-s-federal-government-tax-revenue-3305762

long ascendancy indicates sustained growth, punctuated by occasional deviations. The resilience demonstrated in revenue generation during this period underscores the adaptability of the U.S. fiscal system and its capacity to recover from economic downturns, showcasing a robust financial foundation.

Significantly, economic downturns, exemplified by the aftermath of the 2008 financial crisis[2], have distinctly influenced tax revenue patterns. The years 2009 and 2010 witnessed a noticeable dip in revenue, mirroring the broader economic challenges faced during that tumultuous period. This dip serves as a poignant reminder of the immediate impact that economic crises can have on government revenue streams. However, what stands out amidst these challenges is the subsequent period of recovery. In the aftermath of the downturn, the subsequent years displayed a gradual but discernible rebound in tax revenue. This recovery underscores the nation's resilience and adaptive capacity, showcasing an ability to learn from economic setbacks and implement strategies that contribute to financial rejuvenation.

These observed trends serve as a powerful illustration of the interconnectedness between economic health and government revenue. The impact of economic downturns on tax revenue highlights the delicate balance between economic stability and fiscal outcomes. The subsequent recovery emphasizes the dynamic nature of the U.S. economy and its ability to respond positively to strategic interventions. These findings underscore the importance of fiscal strategies that are not static but adaptive, capable of navigating the ever-changing economic landscapes. Policymakers and financial planners can draw valuable insights from these trends, recognizing the need for flexibility and resilience in crafting strategies that promote both economic recovery and sustained government revenue growth.

We pay attention to the real medium household income that can further help analyze how it can have an impact on the Tax revenue of a government. The graph [3]below can further help evaluate its effects. The dataset, spanning nearly four decades, reveals overarching patterns of long-term growth interspersed with fluctuations reflective of the dynamic economic landscape. Notably, the early 2000s exhibit cyclical income patterns, leading up to the 2008 financial crisis just like the previous finding, which prompted a period of recovery in the following years. The 2010s, however, are marked by income stagnation, emphasizing the need for a nuanced examination of the factors influencing household prosperity during this timeframe. The dataset culminates with a noteworthy surge in real median household income in 2019, reaching $78,250, suggesting a potential positive economic shift. The subsequent years, encompassing the economic impact of

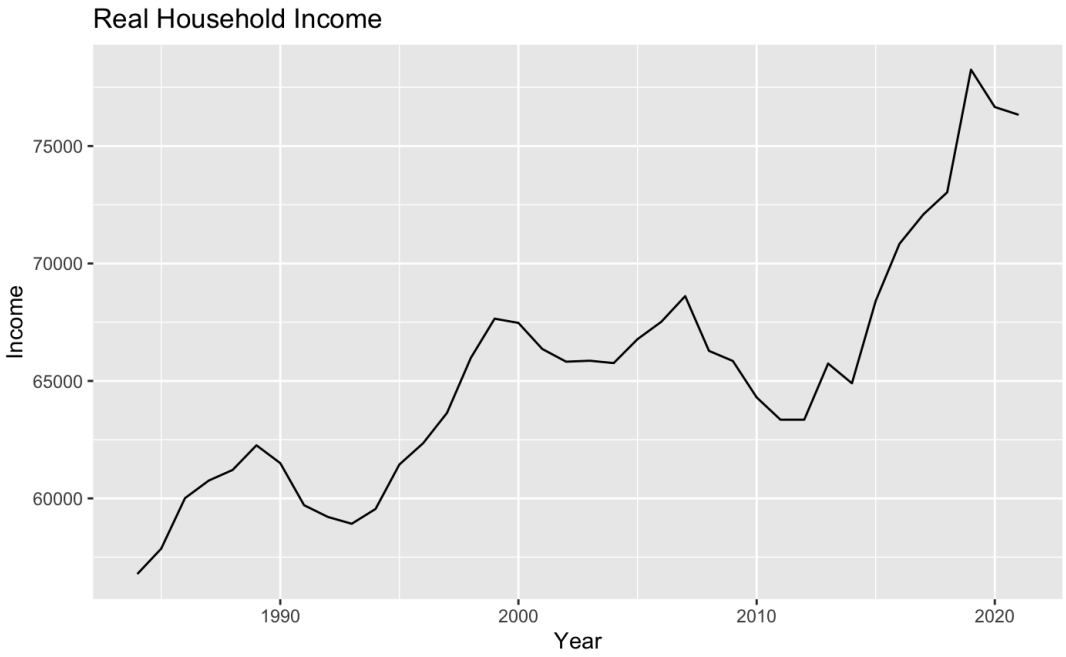
---

[2] "Financial Crisis 2008" Weinberg, J. (n.d.). *The great recession and its aftermath*. Federal Reserve History. https://www.federalreservehistory.org/essays/great-recession-and-its-aftermath

[3] Graph 2 "Real Medium Household Income" *Real median household income in the United States*. FRED. (2023, September 12). https://fred.stlouisfed.org/series/MEHOINUSA672N

the COVID-19 pandemic, demonstrate relative income stability, prompting a deeper exploration into the complex interplay of economic forces shaping household finances.

Graph 2 Real Medium Household Income



Understanding the intricacies of real median household income trends provides valuable context for broader economic discussions. The link between income dynamics and tax revenue is unmistakable, as household incomes directly impact government revenue through taxation. The cyclical patterns and recovery observed post-recession highlight the resilience of households and their role in contributing to the overall economic health.

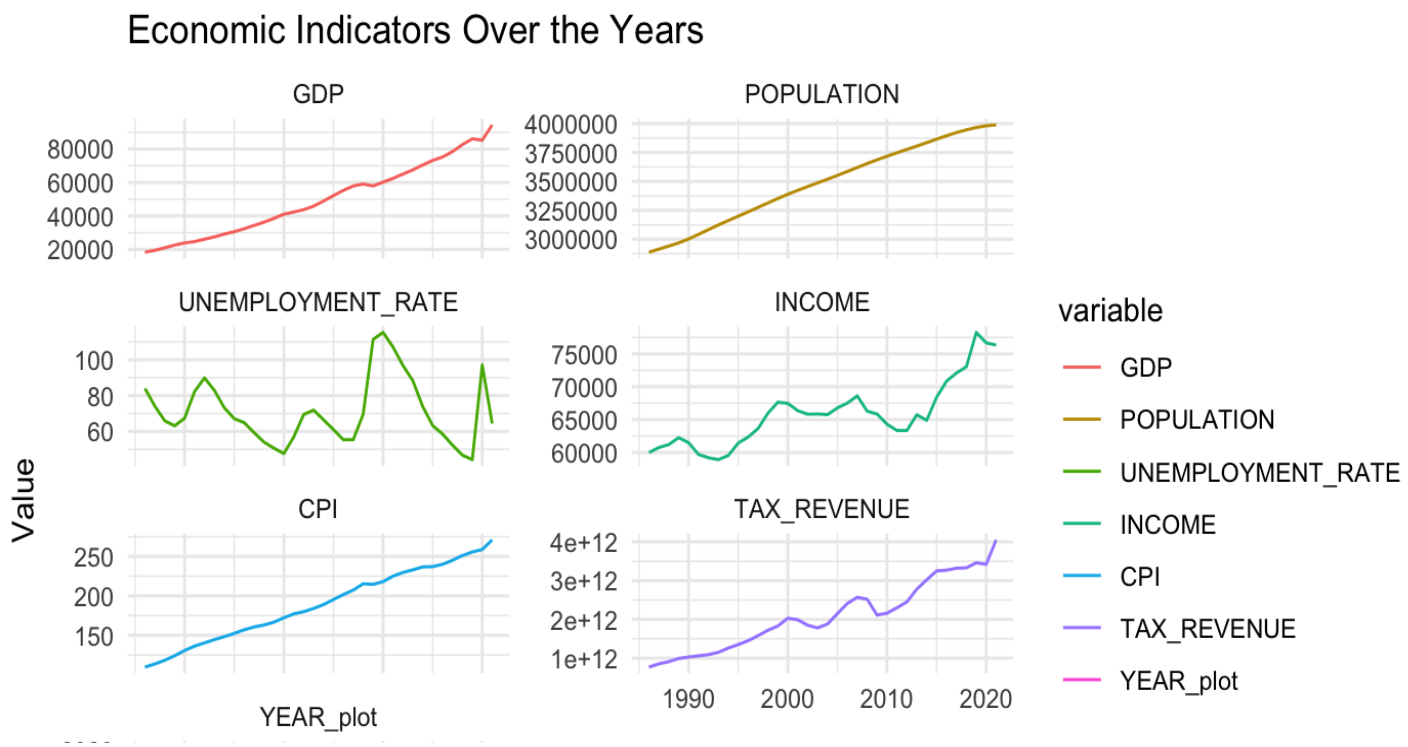
In addition to real household income, this comprehensive overview spans four pivotal economic indicators – Gross Domestic Product (GDP), Population, Unemployment Rate, and Consumer Price Index (CPI) – covering the years 1986 to 2021. These indicators collectively offer nuanced insights into the economic dynamics and overall well-being of the United States during this extensive period.

Gross Domestic Product (GDP): The Gross Domestic Product [4]demonstrates a consistent upward trajectory, reaching $94,376.123 billion in 2021. Cyclical patterns and economic peaks, such as those in 2000, 2007, and 2019, illustrate phases of robust economic activity. The resilience and recovery post-economic downturns underscore the adaptability of the U.S. economy.

---

[4] "Graph 3 Economic Indicators Over the Years" *Gross domestic product*. FRED. (2023a, October 26). https://fred.stlouisfed.org/series/GDP

Graph 3 Economic Indicators Over the Years



The population [5]steadily increased from 288 million in 1986 to 398 million by 2021, showcasing demographic stability and predictability. Incremental yearly growth contributes to the overall expansion of the U.S. population. The size and composition of the population influence economic indicators such as GDP, employment rates, and consumer spending. Unemployment rates exhibit cyclical patterns, influenced by economic events, with peaks during downturns such as the early 1990s, 2009, and 2020. Recovery and resilience are evident, with declining rates post-recessions, reflecting the adaptability of the U.S. labor market. Global events, like the COVID-19 pandemic in 2020, have a profound impact on unemployment rates.

Finally, the Consumer Price Index demonstrates a steady upward trend, indicating a consistent increase in the overall price level of goods and services. Inflationary periods in the late 1980s and early 1990s are followed by periods of relative stability. The CPI experienced fluctuations in 2020-2021, corresponding with the global challenges posed by the COVID-19 pandemic.

---

[5] "Graph 3 Economic Indicators Over the Years" *Population*. FRED. (2023b, October 27). https://fred.stlouisfed.org/series/POPTHM

II. Results

Model 1: In the initial model, the relationship between Gross Domestic Product (GDP) and tax revenue is found to be statistically significant. The coefficient for GDP stands at 3.98*** ($p<0.01$), indicating a positive impact on tax revenue. This implies that for every unit increase in GDP, tax revenue is expected to increase by approximately 3.98 units. This finding underscores the pivotal role of economic output in influencing the nation's tax revenue.

Model 2: Building upon the initial model, the second model reinforces the significant positive relationship between GDP and tax revenue. The coefficient for GDP remains statistically significant at 4.49*** ($p<0.01$), emphasizing the robust impact of economic output on tax revenue. Notably, the population coefficient lacks statistical significance, suggesting that population dynamics do not consistently influence tax revenue in this context.

In Model 3, the positive association between GDP and tax revenue persists, with a coefficient of 3.93*** ($p<0.01$). While the impact of population remains non-significant, the introduction of the unemployment rate reveals a significant negative correlation. The coefficient for the unemployment rate is -6.318*** ($p<0.01$), indicating that an increase in the unemployment rate is associated with a decrease in tax revenue.

Model 4: Continuing the analysis, Model 4 reaffirms the positive correlation between GDP and tax revenue (coefficient = 4.12***, $p<0.01$). Population maintains its non-significant impact on tax revenue. The unemployment rate remains negatively significant (coefficient = -6.80***, $p<0.01$). However, the introduction of median household income shows a negative relationship that is not statistically significant.

Model 5: In the final model, the positive relationship between GDP and tax revenue remains robust (coefficient = 3.11***, $p<0.01$). Population's coefficient indicates a negative relationship, but it lacks statistical significance. The unemployment rate remains negatively significant (coefficient = -7.11***, $p<0.01$). Median household income shows a non-significant negative relationship, and the Consumer Price Index (CPI) is not statistically significant in influencing tax revenue. These comprehensive findings provide valuable insights into the intricate dynamics of the factors impacting tax revenue, with GDP consistently emerging as a key driver.

III. Methods

The integration of FRED and the U.S. Bureau of Labor Statistics datasets forms the basis for an extensive examination of correlations related to tax revenue and its determinants. In total, five linear regression models [6](Model 1-5) are constructed to analyze the impact of key economic indicators on tax revenue. These models, implemented using the R programming language, employ tax revenue as the dependent variable and incorporate independent variables such as Gross Domestic Product (GDP), population, unemployment rate, median household income, and

[6] Model 1, “Model 1: All Linear Regressions Results,” Generated by R

the Consumer Price Index (CPI). The data undergoes meticulous preprocessing, involving cleaning procedures to handle missing values and outliers, and feature selection to focus on
Model 1: All Linear Regressions Results

relevant variables. Exploratory Data Analysis (EDA) employs graphical representations and temporal analysis to elucidate trends and relationships within the dataset, with a particular focus on key economic indicators. The regression analysis assesses the statistical significance of each variable, interpreting coefficients to discern the strength and direction of their impact on tax revenue. Model validation techniques, including cross-validation and diagnostic checks, ensure the reliability and generalizability of the constructed models. The results provide valuable insights into the intricate dynamics of factors influencing tax revenue, with GDP consistently emerging as a key driver. This comprehensive approach aims to contribute meaningful findings to the understanding of the complex relationship between economic and demographic indicators and tax revenue.

IV. Summary

This project endeavors to develop a predictive model for annual tax revenue in the United States, utilizing a dataset spanning from 1984 to 2021 that includes key economic indicators such as Gross Domestic Product (GDP), population, unemployment rate, median household income, and the Consumer Price Index (CPI). The overarching objective is to offer policymakers and stakeholders a robust tool for anticipating tax revenue trends, emphasizing the intricate interplay between economic shifts and fiscal policy adjustments. The methodology encompasses data cleaning, exploratory data analysis (EDA), and the construction of five linear regression models using the R programming language. Through a meticulous analysis, the project highlights the resilience of the U.S. tax system, exemplified by a decade-long ascendancy in tax revenue from 2010 to 2019 despite occasional deviations. Economic downturns, exemplified by the aftermath of the 2008 financial crisis, are explored, revealing the immediate impact on tax revenue followed by a gradual rebound, emphasizing the adaptability and dynamic nature of the U.S. fiscal system. The relationship between economic health, government revenue, and the intricate dynamics of household income are thoroughly examined, providing valuable insights into the interconnectedness of these factors and their influence on tax revenue.

The results of the regression models underscore the pivotal role of GDP in shaping tax revenue, with consistent positive correlations. Population dynamics, unemployment rates, and median household income also emerge as significant factors, showcasing their nuanced impact on tax revenue fluctuations. The project concludes with actionable recommendations for policymakers, emphasizing the need for adaptive fiscal strategies that recognize the dynamic nature of the U.S. economy. By contributing valuable insights into the complex relationship between economic and demographic indicators and tax revenue, this project aims to enhance fiscal planning and decision-making processes for a more resilient and responsive fiscal landscape.

Model 1: All Linear Regressions Results

Results

	Dependent variable:				
	TAX REVENUE				
	(1)	(2)	(3)	(4)	(5)
GDP	3.98***	4.49***	3.93***	4.12***	3.11***
	(1.194)	(6.63)	(4.56)	(5.68)	(9.66)
POPULATION		-3.21	2.28	-3.00	-8.36
		(4.15)	(2.85)	(3.02)	(6.93)
UNEMPLOYMENT_RATE			-6.318***	-6.80***	-7.11***
			(9.93)	(1.307)	(1.31)
INCOME				-5.64	-3.51
				(9.81)	(1.05)
CPI					1.03
					(8.02)
Constant	1.06	9.73	4.99	9.93	2.02
	(6.47)	(1.12)	(7.61)	(1.15)	(1.39)
Observations	36	36	36	36	36
R^2	0.970	0.971	0.987	0.987	0.988
Adjusted R^2	0.970	0.969	0.986	0.986	0.986
Residual Std. Error	1.55 (df = 34)	1.55(df = 33)	1.05 (df = 32)	1.06(df = 31)	1.05 (df = 30)
F Statistic	1.11*** (df = 1; 34)	5.50*** (df = 2; 33)	8.19*** (df = 3; 32)	6.01*** (df = 4; 31)	4.91*** (df = 5; 30)
Note:					*p<0.1; **p<0.05; ***p<0.01

VII. References

- U.S Tax Revenue Over Years 1963-2021" Amadeo, K. (2022, December 12). *U.S. Federal Government Tax Revenue*. The Balance. https://www.thebalancemoney.com/current-u-s-federal-government-tax-revenue-3305762
- "Financial Crisis 2008" Weinberg, J. (n.d.). *The great recession and its aftermath*. Federal Reserve History. https://www.federalreservehistory.org/essays/great-recession-and-its-aftermath
- Graph 2 "Real Medium Household Income" *Real median household income in the United States*. FRED. (2023, September 12). https://fred.stlouisfed.org/series/MEHOINUSA672N
- Graph 3 Economic Indicators Over the Years" *Gross domestic product*. FRED. (2023a, October 26). https://fred.stlouisfed.org/series/GDP
- Graph 3 Economic Indicators Over the Years" *Population*. FRED. (2023b, October 27). https://fred.stlouisfed.org/series/POPTHM
- [1] Model 1, "Model 1: All Linear Regressions Results," Generated by R
- https://fiscaldata.treasury.gov/
- https://www.thebalancemoney.com/current-u-s-federal-government-tax-revenue-3305762