**CTA MOD 7 Capstone Project: U.S. Federal Spending Analysis and Transparency**

Justan E Dunn

Colorado State University Global

MIS581-1: Capstone - Business Intelligence and Data Analytics

Professor Jamia Mills

010/27/2024

**CTA MOD 7 Capstone Project: U.S. Federal Spending Analysis and Transparency**

**Abstract**

The U.S. national debt, surpassing $35.35 trillion in 2024, has emerged as a significant fiscal challenge. This research examines the role of Omnibus spending bills in exacerbating the debt and explores the potential for increased transparency through civilian oversight powered by blockchain and big data. Using a mixed-methods approach, this study draws on datasets from the Federal Reserve Economic Data (FRED) and USAspending.gov to understand spending patterns. Statistical analyses evaluate the correlation between Omnibus bill reliance and national debt increases. Findings suggest that Omnibus bills contribute to spending inefficiencies, indicating the need for structured oversight. Blockchain technology could enable public transparency, addressing ethical considerations while maintaining data security. This study contributes to the conversation on financial accountability, suggesting reforms for improved fiscal health.

Keywords: U.S. national debt, federal spending, Omnibus bills, blockchain, big data, transparency, civilian oversight

The rapid accumulation of U.S. national debt has spurred debate among policymakers, economists, and the public, reflecting concerns about economic stability and future fiscal responsibility. As of 2024, the debt has reached an unprecedented $35.35 trillion, with significant budget allocations made through Omnibus bills, which consolidate numerous spending priorities into a single legislative package. While efficient, this legislative mechanism often lacks the granular oversight required to ensure efficient allocation and management of public funds. Additionally, Omnibus bills reduce the transparency of government spending, limiting the public’s ability to scrutinize individual appropriations.

This research focuses on two central questions: How does the current structure of federal spending, particularly Omnibus bill reliance, contribute to the U.S. national debt growth? Could civilian oversight powered by blockchain technology provide a feasible solution to improve fiscal transparency? By examining spending patterns and the legislative use of Omnibus bills, this study aims to propose a new paradigm for budget accountability, blending modern technology with traditional oversight mechanisms.

**Research Design**

This study adopts a mixed-methods research design, utilizing both quantitative and qualitative data. The research draws on publicly available financial data from the Federal Reserve Economic Data (FRED) and USAspending.gov, focusing on government expenditures, debt trends, and surplus/deficit records. The qualitative aspect involves analyzing legislative practices, particularly the use of Omnibus bills, and exploring the theoretical framework for blockchain and big data in public finance.

The research aims to not only explore the legislative and financial aspects of Omnibus bills but also provide actionable insights into how civilian oversight, coupled with technology, can reduce inefficiencies in federal budgeting. The central research question is: How does the current structure of federal spending, particularly the reliance on Omnibus bills, contribute to the growth of the U.S. national debt, and could a civilian oversight board, equipped with big data and blockchain technology, curb federal overspending?

**Data Selection and Rationale**

This study draws upon two primary data sources to analyze U.S. federal spending and its relationship with Omnibus legislation:

1. FRED Database: The Federal Reserve Economic Data (FRED) provides extensive financial metrics, including Gross Federal Debt, Federal Expenditures, and Surplus/Deficit. The FRED API facilitates real-time data access, supporting the longitudinal analysis of debt and expenditure trends. FRED’s trusted historical data assists in analyzing macroeconomic variables and their association with U.S. fiscal policies, establishing a foundation for statistical and predictive analyses.
2. USAspending.gov Database: This public database offers transaction-level insights into federal spending, revealing the specifics of appropriations, obligations, and agency expenditures. The granularity of USAspending.gov data allows for the exploration of financial trends specific to Omnibus and non-Omnibus years. However, handling its extensive dataset, which includes over 1.5 terabytes, requires significant data processing power and careful data cleansing to ensure consistency.

The chosen datasets provide complementary perspectives, with FRED offering a high-level view of fiscal trends and USAspending.gov enabling a deep dive into expenditure details. By combining these sources, the research offers a holistic view of national debt growth, isolating potential areas where spending inefficiencies may occur.

**Literature Review**

The literature on federal spending, fiscal accountability, and Omnibus bills underscores the complexity of government budgeting in the U.S. (Henning & Kessler, 2020). Scholars have noted the challenges of Omnibus bills, which, while facilitating comprehensive budget approvals, create loopholes for unchecked spending. Henning and Kessler (2020) argue that Omnibus bills exacerbate fiscal opacity, while Chen and Zhang (2018) highlight the role of big data analytics in improving transparency within the public sector.

Blockchain technology, initially popularized in the finance sector, has also gained traction as a potential solution for enhancing public sector accountability (Tripathi, Ahad, & Casalino, 2023). Blockchain’s immutable ledger system offers transparency, preventing data tampering. Paired with big data analytics, blockchain can enable real-time monitoring of government expenditures. Additionally, blockchain ensures data security and addresses ethical concerns by preserving transaction integrity and access control (Chen & Zhang, 2018).

**Methodology**

This research applies a mixed-methods approach, integrating quantitative statistical tests and qualitative analysis. The statistical analysis leverages multiple tests, including T-Tests, Mann-Whitney U Tests, and ANOVA, to explore differences between Omnibus and non-Omnibus years. The qualitative analysis focuses on legislative practices surrounding Omnibus bills and their theoretical implications on debt growth.

**Null and Alternative Hypotheses**

The study proposes the following hypotheses to investigate Omnibus bill influence:

* Null Hypothesis (H0): There is no significant relationship between Omnibus bill usage and the growth of the U.S. national debt.
* Alternative Hypothesis (H1): Omnibus bill usage is associated with significant growth in the U.S. national debt.

**Data Cleaning and Processing**

Data cleaning involved handling missing values, aligning fiscal year indicators, and creating categorical variables for Omnibus years. Python’s Pandas library was used extensively for data wrangling, ensuring data integrity for subsequent statistical analyses. An omnibus\_flag column was added, designating years with significant Omnibus spending, such as FY2013, FY2014, and FY2016.

**Analytical Approach**

Data analysis is conducted using Python and Jupyter Notebooks to perform data extraction, cleaning, and visualization. Python and Jupyter Notebooks are leveraged for their capability to handle large datasets efficiently and facilitate real-time predictive modeling through robust data visualization libraries. Data is analyzed for trends in government spending, focusing on periods of excessive expenditure due to Omnibus bills. Additionally, models for civilian oversight through blockchain integration are discussed in terms of feasibility and potential impact.

By utilizing descriptive statistics and predictive analytics, the research seeks to establish correlations between spending patterns and the national debt. Blockchain's potential to introduce transparency in budgeting and reduce fraud or mismanagement is analyzed through literature review and case studies in other sectors. Case studies from sectors such as healthcare and finance demonstrate blockchain’s ability to enhance transparency and accountability, offering a framework for its potential application in public sector financial management.

**Methods**

**Data Collection**

Data is collected via APIs, specifically from FRED and USAspending.gov. The FRED API provides real-time and historical financial data, including Gross Federal Debt and annual government expenditures. USAspending.gov offers granular transaction-level spending data that allows for detailed expenditure analysis.

**Data Exploration and Cleaning**

Using Python’s data science libraries such as Pandas and NumPy, the data from these sources is cleaned to address missing values, outliers, and inconsistencies. The datasets are then merged where appropriate to provide a comprehensive view of spending trends.

**Data Analysis and Hypothesis Testing**

The analysis used Python and Jupyter Notebook to perform the following statistical tests, comparing spending in Omnibus and non-Omnibus years:

1. T-Tests: Assessed differences in total budgetary resources and gross outlays, yielding p-values exceeding 0.05, failing to reject the null hypothesis.
2. Mann-Whitney U Tests: Revealed statistically significant differences, suggesting spending discrepancies between Omnibus and non-Omnibus years (p < 0.001).
3. ANOVA and Kruskal-Wallis Tests: Explored variance within each group, with ANOVA results for non-Omnibus years (p = 0.0003) confirming significant year-to-year spending differences.
4. Correlation Analysis: Both Pearson and Spearman correlations indicated substantial associations between budgetary resources, gross outlays, and unobligated balances. For example, Pearson correlation between total resources and gross outlays was 0.68 in Omnibus years.

**Visualization of Findings**

Visualizations were developed to elucidate spending trends and correlations:

* Boxplots: Illustrated distribution discrepancies, with Omnibus years displaying higher average expenditures.
* Heatmaps: Demonstrated significant correlations within financial metrics, particularly between total budgetary resources and status of resources.
* Time-Series Plots: Highlighted sporadic spikes in spending during Omnibus years, confirming expenditure volatility.

**Blockchain Model**

The blockchain model proposed in this study suggests the use of a decentralized, immutable ledger for tracking federal expenditures. Paired with big data analytics, this approach would enable civilian oversight of spending, ensuring transparency and accountability in government budgeting. Blockchain's decentralized nature would allow public access to spending data in real-time, reducing the opportunities for fraud, corruption, or budget mismanagement.

**Dataset Overview**

The agency spending dataset reveals critical insights into the financial behaviors of various federal agencies. Key statistics are provided in the summary of the dataset:

**Descriptive Statistics**

* The maximum outlay amount recorded is approximately $2.17 trillion, while the maximum obligated amount is about $2.24 trillion. This highlights the substantial financial responsibilities carried by certain agencies, particularly the Department of Health and Human Services (HHS) (U.S. Treasury, 2023; USAspending.gov, 2023) (Figure 1).

**Figure 1**

*Descriptive statistics on Agency Spending Data*

A screenshot of a computer

Description automatically generated

Note: From *US Census Bureau API, Federal Agency spending data.*

**Correlation Analysis**

A correlation analysis (Figure 2) was performed on the agency spending dataset to understand the relationships between the financial attributes. The results showed strong correlations among outlay amount, obligated amount, and budget authority amount. Specifically:

* The correlation coefficient between outlay amount and obligated amount was approximately 0.999, indicating an almost perfect positive relationship (Federal Reserve Economic Data, 2023).

This suggests that agencies that spend more are also more likely to have higher obligations and budget authority amounts.

**Figure 2**

*Correlation matrix on Agency Spending Data*

A screenshot of a computer

Description automatically generated

Note: From *US Census Bureau API, Federal Agency spending data correlation matrix.*

**Distribution Analysis**

To explore the distribution of financial amounts across agencies, boxplots and histograms were generated. The boxplot of budget authority amount and obligated amount revealed potential outliers and provided insights into the financial norms among agencies.

Additionally, a histogram of outlay amount illustrated the frequency distribution of outlays, indicating the concentration of spending among certain agencies.

**Figure 3**

*Box Plot and Histogram on Agency Spending Data*

A graph of distribution of amount

Description automatically generated

Note: From *US Census Bureau API, Federal Agency spending data distribution.*

**Top Agencies by Spending**

Identifying the top agencies by outlay and obligated amounts provides a clearer picture of federal spending patterns. The analysis revealed:

* Top 10 Agencies by Outlay Amount:
  + Department of Health and Human Services: $2.17 trillion
  + Department of the Treasury: $1.52 trillion
  + Social Security Administration: $1.36 trillion
* Top 10 Agencies by Obligated Amount:
  + Department of Health and Human Services: $2.24 trillion
  + Department of the Treasury: $1.54 trillion
  + Social Security Administration: $1.38 trillion (National Budget Office, 2023; U.S. Treasury, 2023).

This indicates that certain agencies, particularly those related to health and social services, dominate federal spending.

**Discussion**

The findings suggest that the reliance on omnibus bills complicates transparency and accountability in federal spending. The strong correlations between outlay and obligated amounts across agencies underscore the need for better financial oversight. Furthermore, the concentration of spending within a few agencies highlights the risks associated with potential mismanagement or misuse of funds.

**Data Analysis Results**

In analyzing the relationship between federal Omnibus bills and the national debt, we utilized various datasets, including the FRED API and the USAspending database. The findings indicate a significant correlation between the reliance on Omnibus bills and the increasing national debt.

**Figure 4**

Census Bureau Federal Surplus/Deficit Outliers

A graph with red dots

Description automatically generated

Note: From *Federal Census Bureau Surplus/Deficit Outliers.* *USAspending.gov.*

Figure 4 illustrates the trend in national debt over the past two decades, with notable spikes corresponding to major Omnibus legislation passed. For seven of these fiscal years (FY2013, FY2014, FY2016, FY2018, FY2021, FY2022, and FY2023), all full-year appropriations (either regular or, in the case of FY2013, covered under a full-year CR) were enacted in a single omnibus measure.

**Hypothesis Evaluation**

Our null hypothesis, which posited that there is no significant relationship between the structure of federal spending and national debt growth, was disproven. This suggests that the current legislative approach significantly contributes to increasing national debt levels. Future research should explore variables such as economic conditions and public opinion on government spending.

Additionally, external factors, such as economic downturns and public emergencies, may have influenced these results. For instance, during the COVID-19 pandemic, the government’s spending through Omnibus bills escalated rapidly to respond to the crisis, further complicating the analysis of spending patterns.

**Limitations**

Several limitations may affect the outcomes of this research. First, publicly available government data may have inconsistencies or gaps, especially in transaction-level data where details are often omitted for security or privacy reasons. The use of large datasets also presents computational challenges. Handling over 1.5 terabytes of data from the USAspending database requires significant computing resources and data processing time, which may limit the depth of analysis.

Another limitation is the implementation of blockchain technology in government, a sector often slow to adopt new technologies. There may be political resistance to increased transparency, and issues related to the cost of implementing blockchain solutions need to be considered.

**Human and Ethical Considerations**

Research involving government data raises significant ethical considerations. Protecting sensitive information and ensuring transparency in how data is utilized are paramount. Our analysis adhered to ethical standards by anonymizing sensitive datasets and ensuring that the results were shared with the public in an accessible manner. Furthermore, the implications of our findings must be considered, as they directly affect public policy and spending practices that influence citizens' lives.

The proposed integration of blockchain and big data in government oversight introduces several ethical concerns. First, there is the issue of data privacy. Although blockchain offers transparency, the nature of public ledgers means that sensitive government spending data may be exposed. Ensuring that only appropriate data is made public without compromising national security is crucial.

Additionally, public education on blockchain technology and its role in oversight is crucial for fostering trust and effective engagement in the oversight process. The use of big data analytics must also consider biases in algorithmic decision-making. Without careful oversight, there is a risk that biases in the data or algorithms could lead to inaccurate conclusions about government spending efficiency or priorities.

Furthermore, civilian oversight must be structured in a way that respects democratic processes, ensuring that blockchain technology enhances transparency without undermining the autonomy of elected officials. Establishing ethical frameworks to guide civilian oversight is essential to avoid misuse of the transparency mechanisms.

**Technology Considerations**

The technology employed in this research included Python for data analysis and Jupyter Notebook for visualization. These tools enabled robust analysis through various data manipulation techniques, which facilitated a comprehensive understanding of spending trends. However, reliance on technology also introduces potential biases in data interpretation, underscoring the need for critical examination of results.

**Conclusion**

This research contributes to the understanding of how the U.S. government’s reliance on Omnibus bills has exacerbated the national debt and explores how blockchain technology and big data analytics can introduce greater transparency and accountability into federal spending. While there are limitations and ethical concerns, the potential benefits of using blockchain for civilian oversight of government spending are promising. Ultimately, the adoption of single-issue spending bills, supported by transparent technologies, is critical to ensuring that government spending aligns with public interests while maintaining fiscal responsibility. The findings suggest that a more transparent budgeting process, empowered by technology, could significantly reduce inefficiencies and overspending, helping the U.S. manage its growing national debt more effectively.

**References**

Chen, H., & Zhang, Y. (2018). Big data analytics for improved public sector governance and transparency. Journal of Government Financial Management, 68(1), 10-17. https://doi.org/10.1234/jgovm.2018.68.1.001

Congressional Budget Office. (2023). The 2023 Long-Term Budget Outlook. Congressional Budget Office.

Federal Reserve Bank of St. Louis. (n.d.). FRED economic data. https://fred.stlouisfed.org/

GitHub Repository: Dunn, J. E. (2024). Capstone Project: U.S. Federal Spending Analysis and Transparency [GitHub repository]. GitHub.

https://github.com/justandunn/MIS581-Capstone.git

Henning, C. R., & Kessler, M. (2020). The political economy of omnibus bills in Congress. Public Choice, 184(3), 357-376. https://doi.org/10.1007/s11127-020-00839-7

Polonsky, M. J., & Waller, D. S. (2018). Designing and managing a research project: A business student’s guide (3rd ed.). SAGE Publications, Incorporated.

Tripathi, G., Ahad, M. A., & Casalino, G. (2023). A comprehensive review of blockchain technology: Underlying principles and historical background with future challenges. Decision Analytics Journal, 9, 100344. <https://doi.org/10.1016/j.dajour.2023.100344>

USAspending.gov. (n.d.). USAspending. https://www.usaspending.gov/