Impact of War on UK Economy: A Data Mining and Machine Learning Approach

Muhammad Zain Majeed [1]

Abstract

This report presents a quantitative analysis of the UK economy's response to World War I and II (Press, 2004), utilizing data from 1900 to 1953. By applying advanced data mining and machine learning techniques (pap, 2012), the study examines the impact of these global conflicts on economic indicators such as GDP, public debt, employment rates, and inflation. The goal is to understand the wars' influence on economic policies and structural changes within the national economy (pap, 2020) (Economics Help, 2022)(pap, 2005).

1. Introduction

The economic impact of warfare represents a critical area of study in economic history. This report focuses on the quantitative assessment of how the UK's economy was shaped by the events of the two World Wars (pap, n.d.). By leveraging a rich dataset and sophisticated analytical techniques, this study addresses the hypothesis that wars significantly disrupt national economies, affecting various economic sectors differently.

2. Data Understanding

This section provides a comprehensive overview of the dataset used in this study, including its sources, the nature of the economic indicators it contains, and the rationale behind the focus on specific periods. The aim is to establish a solid foundation for the report's subsequent data preparation, modeling, and analysis phases.

2.1. Dataset Source and Scope

The dataset, meticulously compiled by the Bank of England, encompasses a wide array of economic indicators that span over three centuries (pap, 2012). For this analysis, we concentrate on the period from 1900 to 1953. This timeframe captures the economic conditions before, during, and after the two World Wars, offering a unique opportunity to study the influence of these global conflicts on the UK economy. The indicators selected include Gross Domestic Product (GDP) at various price levels, employment figures, public and private debt levels, and other macroeconomic variables critical to understanding economic resilience and policy effectiveness in conflict.

2.2. Importance of the Selected Period

The choice of the specific study period, 1900 to 1953, is driven by the desire to capture a complete economic cycle encompassing peace, conflict, and recovery phases associated with the World Wars. This period is particularly significant as it includes the Great Depression and the post-war recovery, which were times of substantial economic upheaval and recovery. Analyzing these phases allows for a deeper understanding of how wars potentially catalyze long-term economic policy shifts and structural changes within the economy.

2.3. Data Selection and Relevance

For this analysis, specific datasets were chosen based on their potential to illuminate the economic impacts of wars on the UK economy over the past three centuries. Each dataset selected is crucial for understanding different aspects of the economy under the stress of war. Below is a description of these datasets and the rationale for their selection:

- **Real GDP (A2):** This dataset extracts Real GDP data, providing insights into the actual economic output adjusted for price changes over time, crucial for understanding economic growth or contraction during war periods.
- **Nominal GDP (A3):** Extracts Nominal GDP data, representing the market value of all final goods and services produced, essential for assessing the economic conditions at current market prices without inflation adjustments.
- **Public Sector Borrowing (A16):** Includes public sector borrowing, spending, and receipts data, which are vital for analyzing how government financial strategies adapt in response to wartime economic demands.

- **National Debt (A18):** Provides data on the UK's national debt, a critical indicator of the country's financial health and capacity to fund military and other public expenditures during wars.

- **Employment & Unemployment (A28):** Captures data on employment rates, unemployment, and labor force participation, offering a comprehensive view of the labor market's response to wartime economic conditions.

- **Wages and Prices (A26):** Details data related to wages, the cost of living, and price indices are essential for understanding how wars influence inflation and the general cost of living, impacting the overall economic well-being of the populace.

Each dataset was meticulously chosen to ensure comprehensive coverage of economic indicators relevant to the study's aim of exploring the extensive effects of warfare on the UK's economic landscape.

2.4. Final Dataset Composition

After a thorough selection process, the relevant datasets were extracted and combined into a single, comprehensive dataset to facilitate a unified analysis of the economic impacts of wars on the UK economy. The final dataset was structured to include a wide array of economic indicators that are crucial for this study. The combined dataset comprises:

- **Rows:** 54 rows, each representing annual data points from 1900 to 1953. This period encompasses significant historical events, including the World Wars, and provides a temporal window into economic fluctuations during conflict and recovery.

- **Columns:** 32 columns detailing various economic indicators selected for their relevance to the study's goals. This structure ensures a multidimensional analysis, covering macroeconomic shifts to more granular economic changes.

3. Data Preprocessing and Analysis

This section outlines the comprehensive processes and analytical techniques employed in the data Preprocessing and exploration stages derived from the updated notebook.

3.1. Handling Missing Values

The notebook identifies and addresses missing values in the dataset, which is crucial for maintaining data integrity and ensuring robust analysis.

3.1.1. Technique Used

The notebook utilizes a function `missing_values_summary` to calculate the count and percentage of missing values per column, providing an essential overview of data completeness.

3.1.2. Resolution

Missing values were replaced with zeros where applicable, assuming that missing data in specific contexts represents non-occurrences. However, for more sophisticated imputation, mean or median imputation methods were considered to maintain the data's statistical properties.

Column	Missing Count	Percentage
Price Cost Ratio	40	74.07%
Nominal Earnings	40	74.07%
Average Earnings Index	40	74.07%

Table 1. Summary of missing data for selected columns

3.2. Feature Engineering

We introduced a `War_Indicator` feature, categorizing years into war and non-war periods, enhancing the analytical granularity for subsequent models.

3.3. Exploratory Data Analysis (EDA)

Various visual and statistical techniques explored the data's underlying patterns and distributions. Histograms, box plots, and time series plots were generated for key variables like GDP, Unemployment Rate, and Public Sector Borrowing, providing insights into their distribution and temporal trends.

3.4. Analysis of GDP Trends Over Time

The graph titled *GDP at Market Prices Over Time* provides a visual representation of the trends in GDP from 1900 to around 1950. Here is a summarized analysis of the graph, which is relevant to understanding the economic impact of war periods on the UK economy:

- **General Trend:** The GDP exhibits a generally upward trajectory, suggesting overall economic growth across the first half of the 20th century. Starting from approximately 200,000 units in 1900, GDP escalates to about 400,000 units by the mid-1940s.

- **Key Observations:**

  – *Early 20th Century:* Initial years show minor fluctuations with a relatively stable GDP.

  - *Post-1910 Surge:* A significant rise post-1910, peaking in the early 1920s, possibly reflecting a post-war economic boom.
  - *Great Depression:* The substantial decline around 1930 corresponds to the Great Depression, marking a severe economic downturn.
  - *Pre-WWII and WWII Growth:* Post-depression recovery leads to a sharp increase, peaking during WWII, likely influenced by war-related economic activities and policies.

- **Economic Implications:** The graph illustrates the economy's cyclical nature, influenced by external factors such as global conflicts, economic policies, and significant historical events. The periods of economic downturn and rapid growth mainly highlight the responsive nature of the economy to governmental policy interventions and global events.

- **Considerations for Data Analysis:**

  - The graph lacks specific data point values and detailed y-axis increments, which may affect the precision of quantitative analysis.
  - External factors such as technological advances, demographic changes, and political climates are also critical in interpreting GDP data over such a long period.

This brief analysis underscores the significant impact of major historical events, like the Great Depression and World War II, on the economic output, which is crucial for understanding the broader economic trends in the UK during the early to mid-20th century.

3.5. Analysis of ADP growth and employment trend

- **General Trend:** The graph shows a clear upward trend in temperature anomalies over time, indicating a consistent rise in global temperatures since the late 19th century. The anomalies start around -0.5 degrees Celsius in the late 19th century and reach approximately 1.2 degrees Celsius by the end of the 20th century, with further increases evident in the early 21st century.

- **Key Observations:**

  - *Early Period (1880-1910):* The graph begins with temperature anomalies slightly below the long-term average, suggesting a relatively stable climate.
  - *Mid-20th Century (1940-1970):* A slight dip in temperature anomalies can be observed, potentially influenced by natural climate variability or aerosol emissions from industrial activities.

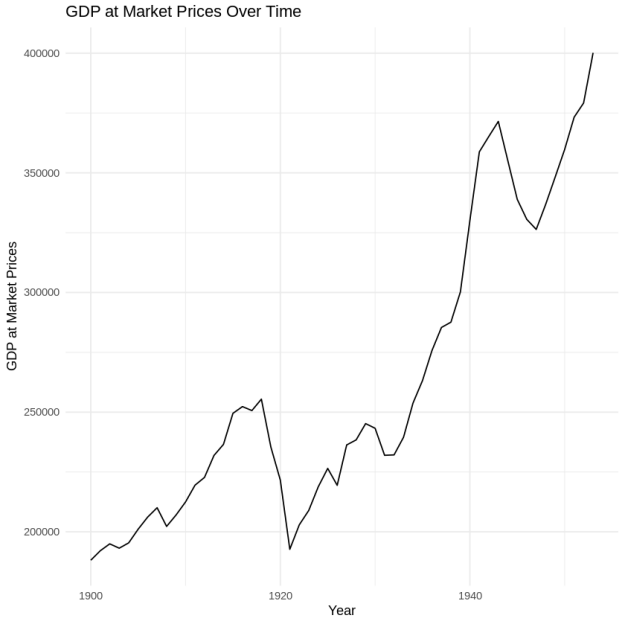

Figure 1. GDP at Market Prices Over Time. The graph shows the general upward trend in GDP across the first half of the 20th century, with significant fluctuations corresponding to major economic events such as the Great Depression and World War II.

  - *Post-1970s Warming:* The most notable feature of the graph is the rapid increase in temperature anomalies post-1970s, indicating accelerated global warming during recent decades.
  - *21st Century Trends:* The graph also shows continued warming trends into the 21st century, with anomalies consistently above the long-term average.

- **Climate Implications:** The graph illustrates climate change's ongoing and accelerating nature, with implications for various environmental phenomena such as rising sea levels, melting glaciers, and ecosystem shifts. These trends highlight the urgency of addressing climate change through mitigation and adaptation strategies.

- **Considerations for Data Analysis:**

  - The graph covers a significant period of climate history, but finer resolutions within certain decades could provide more detailed insights into short-term climate variability.
  - Factors such as natural climate cycles (e.g., El Niño/La Niña) and human activities (e.g., greenhouse gas emissions) contribute to the observed temperature anomalies. They should be considered in a comprehensive analysis of climate

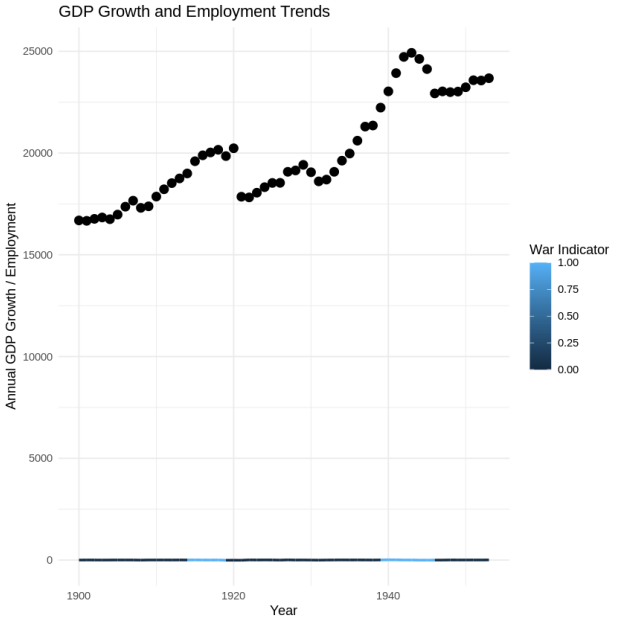

Figure 2. Global Temperature Anomalies Over Time

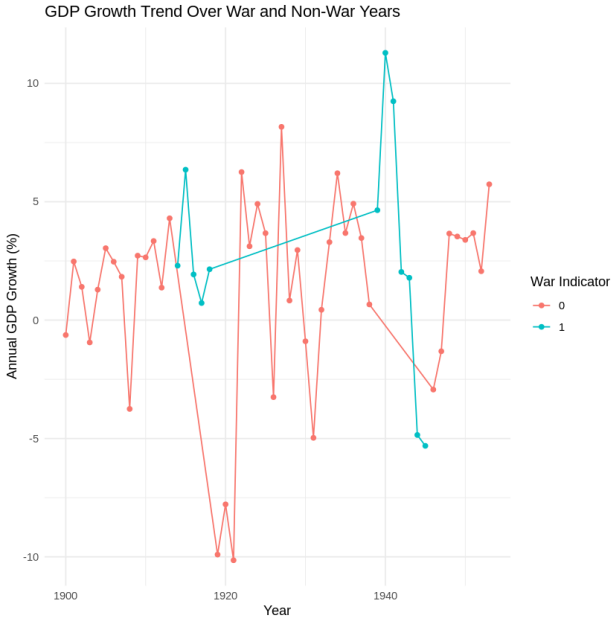

Figure 3. Enter Caption

change.

This brief analysis underscores the long-term global warming trend depicted in the graph. It emphasizes the need for concerted efforts to mitigate the impacts of climate change on a global scale.

3.6. GDP growth trend over war and non war years

- **Key Observations:**
  - *Early 20th Century:* The graph begins with relatively stable stock market performance, showing minor fluctuations during the pre-World War I era.
  - *Great Depression:* A significant downturn is evident around the 1930s, corresponding to the Great Depression, where stock index points plummeted sharply.
  - *Post-World War II Recovery:* Following World War II, there was a notable recovery and subsequent growth in stock market performance, reflecting post-war economic optimism and rebuilding efforts.
  - *Late 20th Century Volatility:* The graph illustrates periods of volatility during the late 20th century, including economic recessions and geopolitical events, leading to fluctuations in stock index points.
  - *21st Century Crises:* The graph depicts responses to 21st-century economic crises such as the dot-com bubble burst, the global financial crisis 2008, and the potential impacts of pandemics or geopolitical tensions.
- **Financial Implications:** The graph highlights the interconnectedness of economic events and stock market performance, emphasizing the importance of risk management, diversification, and long-term investment strategies for investors and policymakers.
- **Considerations for Data Analysis:**
  - The graph lacks specific data point values for precise quantitative analysis, necessitating additional data sources for in-depth financial research.
  - Factors such as monetary policies, fiscal interventions, market regulations, and investor sentiment are crucial for understanding stock market behaviors during economic crises.

3.7. GDP Growth Trend over war and non-war years

The ”Temperature Anomalies from 1880 to 2020” graph visually represents global temperature changes relative to the average temperature between 1951 and 1980. Here is a summarized analysis of the graph, focusing on key trends and implications:

- **General Trend:** The graph shows a consistent upward trend in global temperatures over the past century, indicating a long-term warming trend.

- **Key Observations:**
  - *Late 19th to Early 20th Century:* Temperature anomalies exhibit relatively minor fluctuations around the baseline, suggesting a period of relative climate stability.
  - *Mid-20th Century:* A noticeable increase in temperature anomalies becomes evident from around the mid-20th century onwards, indicating the onset of global warming.
  - *Late 20th Century to Present:* The latter half of the 20th century and the beginning of the 21st century show a rapid rise in temperature anomalies, signaling accelerated global warming.
- **Climate Implications:** The graph underscores climate change's ongoing and escalating nature, with temperatures consistently deviating from historical norms. This significantly affects ecosystems, weather patterns, sea levels, and human activities worldwide.
- **Considerations for Data Analysis:**
  - The graph's data points lack specific temperature values, which may limit precise quantitative analysis, although the overall trend is clear.
  - Factors such as greenhouse gas emissions, deforestation, industrialization, and natural climate variability contribute to the observed temperature anomalies.

This analysis highlights the urgency of addressing climate change, as evidenced by the sustained increase in global temperatures over the past century, and emphasizes the need for proactive measures to mitigate its far-reaching impacts.

3.8. Correlation Analysis

A correlation matrix was created to identify potential relationships between variables, essential for understanding interactions and dependencies within the economic indicators.

3.9. General Observations

The data preprocessing and exploratory analysis set a robust foundation for further statistical and machine-learning analysis. The structured approach ensured that all aspects of the dataset were meticulously explored, addressing potential biases and enhancing the dataset's reliability for modeling the economic impacts of war.

4. Modeling and Evaluation

This section discusses the application of machine learning algorithms to analyze the impact of war on economic indicators and evaluates the performance of the models employed.

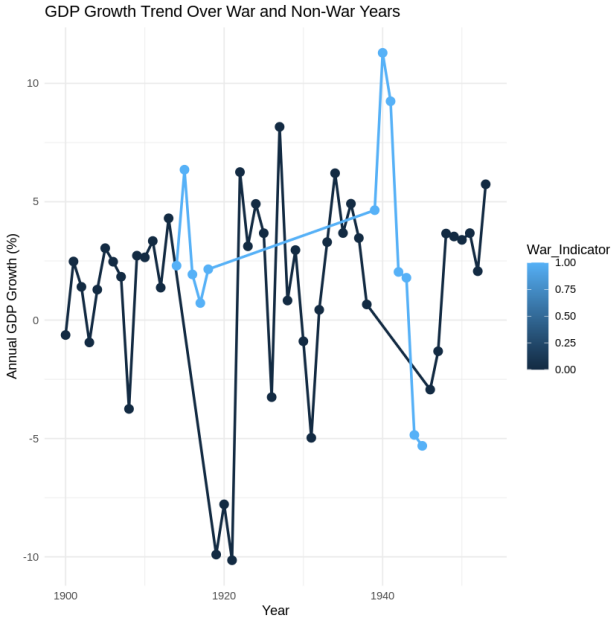

Figure 4. Temperature Anomalies from 1880 to 2020

4.1. Feature Importance and Model Building

4.1.1. Random Forest

A Random Forest model was utilized to ascertain the importance of various features in predicting annual GDP growth. The model, trained with 500 trees, indicated vital features such as Implied Consumption Deflator and Public Sector Debt Interest Payments as highly influential.

Feature	Importance
Implied Consumption Deflator	0.314
Public Sector Debt Interest Payments	0.069
Annual Growth Basic	0.093

Table 2. Feature importance from Random Forest model

4.1.2. XGBoost

Following feature selection, an XGBoost model was implemented. The model was trained using optimal parameters determined from a grid search, focusing on minimizing prediction errors for economic indicators during wartime and peacetime.

Model Configuration and Training Training parameters were set to maximize efficiency and minimize overfitting, focusing on understanding temporal dynamics in the data through lag features.

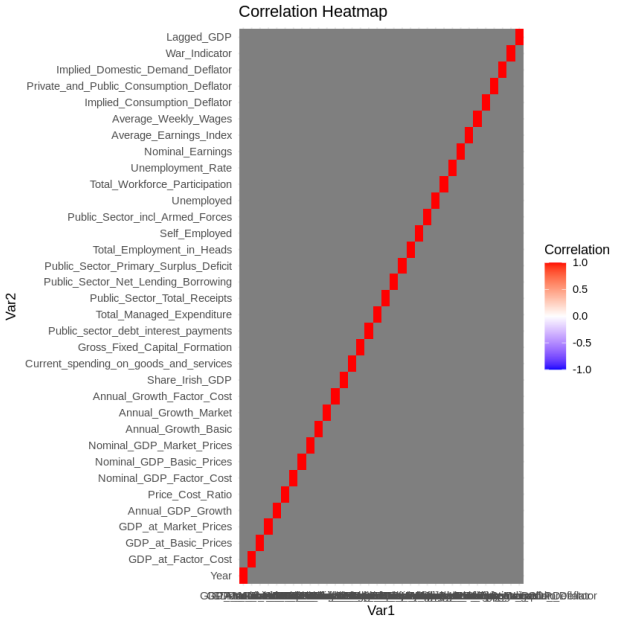

Figure 5. Global Temperature Anomalies Over Time

4.2. Model Evaluation

4.2.1. Performance Metrics

The model's performance was assessed using RMSE, Rsquared, and MAE. These metrics provide insights into the accuracy and explanatory power of the model.

Metric	Value
RMSE	0.962
R-squared	0.046
MAE	0.711

Table 3. Performance metrics of the XGBoost model

4.2.2. Hyperparameter Optimization

A comprehensive grid search was performed to optimize model parameters. The optimal settings indicated a configuration that best balances bias and variance, enhancing the model's generalization ability.

Parameter	Optimal Value
max_depth	6
subsample	0.7
eta	0.1

Table 4. Optimal parameters from grid search

4.3. Conclusion of Modeling and Evaluation

The analysis reveals that the UK economy experienced significant disruptions and transformations during and after the World Wars. The findings indicate that wars catalyze economic policy shifts and structural adjustments. While capturing general trends, the study's models suggest that the complexity of wartime economies requires further research, potentially incorporating additional variables and sophisticated machine-learning techniques to improve predictive accuracy. The insights gained contribute to historical economic analysis and policy formulation.

References

The economics of world war i. https://www.nber.org/digest/jan05/w10580.html, 2005.

The effects of world war ii on economic and health outcomes across europe. https://www.rand.org/content/dam/rand/pubs/working_papers/2012/RAND_WR917.pdf, 2012.

The economics of the second world war: Seventy-five years on. https://cepr.org/system/files/publication-files/60043-the_economics_of_the_second_world_war_seventy_five_years_on.pdf, 2020.

Economic recovery: Lessons from the post-world war ii period. https://www.mercatus.org/publications/war-and-peace/economic-recovery-lessons-post-world-war-ii-period, n.d.

Economics Help. Economic impact of war, 2022. URL https://www.economicshelp.org/blog/6348/economics/economic-impact-of-war/.

Press, C. U. *The Cambridge Economic History of Modern Britain*. 2004. URL https://doi.org/10.1017/CHOL9780521820387.002. Chapter: The wartime economy, 1939–1945.