Bayesian Insights into Aerial Bombing Strategies: An Ordered Logistic Regression Analysis of WWII Target Prioritization Against Germany*

Yunzhao Li

April 18, 2024

During World War II, the strategic decisions about where and how the Allies bombed German industries played an essential role in the Allied campaign. This study employs Bayesian ordered logistic regression to methodically analyze how different factors—such as the industry type, the launching country, the bomb load, and the number of aircraft—shaped these decisions. Focusing on the preference for targeting essential oil refineries, the analysis sheds light on the calculated approach to disrupt Germany's war capabilities effectively. The insights from this research not only clarify the strategic rationale behind the Allies' bombing priorities but also enhance our understanding of their impact on the course of the war.

1 Introduction

World War II reshaped global power dynamics, with the strategic aerial bombing campaigns playing a essential role in determining the outcome of the conflict. The Allied forces' focused bombardment of German industries was pivotal, aiming to dismantle the economic backbone essential for Germany's war efforts. While considerable research, such as Pape's analysis of strategic bombing's effectiveness (Pape 1996), has been conducted on the impact of these bombings, less attention has been given to the systematic selection process behind targeting specific industrial assets. This paper addresses this understudied aspect by exploring the intricate decision-making process that guided the Allied forces in targeting German industries (Overy 2013).

*Code and data are available at: https://github.com/yunzhaol/aerial_bomb_priority.git.

The crux of this research lies in a detailed examination of a sampled subset of over 5,000 missions from extensive wartime records, representing a strategic selection designed to analyze the impact of various factors on target prioritization. These factors include the type of industry targeted, the leading country of the mission, the volume of bombs dropped, and the scale of aircraft involvement. Prior literature, such as studies by Grey (Grey 1999), has often provided broad overviews without delving into the interplay of these specific variables. By applying a Bayesian ordered logistic regression model, this study meticulously quantifies the influence of each factor, offering a clearer picture of the strategic considerations that shaped the Allied bombing strategy.

Our findings indicate a deliberate emphasis on disabling critical oil refineries and other vital industrial infrastructures, which were deemed essential for sustaining the German war machine. The analysis underscores that missions with larger bomb payloads and more aircraft were preferentially deployed against these high-value targets. This targeted approach was strategic, aiming not just at destruction but at crippling Germany's ability to sustain its military operations effectively.

The significance of these insights extends beyond the historical narrative of World War II, offering lessons on the allocation of military resources and strategic target selection that are relevant to modern military strategies. Furthermore, this paper enriches the ongoing academic debate about the strategic efficacy and ethical considerations of aerial bombing in warfare, as discussed in broader terms by scholars like Biddle (Biddle 2002) and Morrow (Morrow 2004).

The paper is organized to provide a clear and detailed examination of the strategic target prioritization during World War II aerial campaigns. Following the introduction in Section 1, Section 2 outlines the data employed in the study, describing the sources of historical military data, the variables selected for analysis, and the rationale behind these choices. Section 3 introduces the Bayesian ordered logistic regression model employed to analyze the relationships between our variables of interest, providing a statistical framework for understanding the strategic decisions behind target prioritization during World War II aerial campaigns. Section 4 presents the findings from the Bayesian ordered logistic regression model, detailing how different factors such as target industry and mission country influenced the prioritization of targets. Section 5 examines these results, discussing the strategic implications of the findings, the limitations of the current study, and suggesting avenues for future research to expand upon the insights gained. Section A shows plots for Posterior predictive check and Diagnostics and their analysis.

1.1 Estimand

This study aims to estimate the strategic priorities that guided the Allied forces in their aerial bombing campaigns against German industries during World War II. By analyzing various factors such as the type of industry targeted, the country leading the mission, the tonnage of bombs dropped, and the number of aircraft involved, this research seeks to understand how these variables influenced the decision-making processes regarding target selection. The core goal is to determine the causal effect of these factors on the prioritization of bombing targets, thereby shedding light on how strategic necessities, resource availability, and tactical choices shaped the wartime operations.

2 Data

2.1 Measurement

The dataset used in this study was sourced from Data World, which hosts extensive records from the Theater History of Operations (THOR) dataset, documenting aerial bombings during World War II. Originally consisting of approximately 178,282 records, the dataset underwent significant preprocessing to align with the specific focus of this research. The initial step involved filtering out records where the target location was not Germany, reducing the dataset to 64,948 entries. These records form the basis of the analysis, having been retained for detailed examination and modeling of strategic bombing patterns and their impacts.

The data for this study was systematically downloaded, cleaned, analyzed, modeled and visualized using R (R Core Team 2023), a extensive statistical programming language. The following packages were used for this study

- **`tidyverse`** (Wickham et al. 2021): To streamline the process of data manipulation and visualization.
- **`ggplot2`** (Wickham 2021): Used for its powerful and flexible capabilities in creating various types of visualizations tailored to the needs of this study.
- **`dplyr`** (Wickham, François, et al. 2021): Employed for its intuitive functions to transform and summarize the complex data sets effectively.
- **`bayesplot`** (Gelman, Gabry, et al. 2021): Utilized for creating graphical posterior predictive checks and diagnostic plots.
- **`rstanarm`** (Team 2021): Facilitated the implementation of Bayesian models, providing a straightforward way to fit regression models using Stan.
- **`janitor`** (Firke 2021): Making it simpler to handle the raw data by cleaning variable names and simplifying data structures.
- **`arrow`** (Apache Arrow 2021): Used for efficiently reading and writing large datasets, enhancing data handling capabilities.
- **`knitr`** (Xie 2021): Employed to dynamically generate reports which integrate R code with its outputs, allowing for seamless inclusion of plots and analysis results in the final document.
- ***Telling Stories with Data*** (Alexander 2023): This book was referenced for its code and methodologies in presenting data and statistical information.

Table 1: Aerial Bombing missions against Germany on the main target industries in WWII

tgt_priority	tgt_industry	country_mission	bomb_tons	aircraft_attack
secondary target	railway infrastructure	USA	42	17
secondary target	railway infrastructure	USA	20	8
target of last resort	unidentified targets	USA	3	1
target of last resort	unidentified targets	USA	3	1
target of last resort	unidentified targets	USA	3	1
secondary target	railway infrastructure	USA	5	3

Table 1 presents the first six rows from the cleansed dataset, focusing on aerial bombing missions conducted by the Allies against German industries during WWII.

2.2 Variable

Our analysis focuses on the following variables, with a specific focus on `target_priority` as the dependent variable:

- `target_priority`: The assigned importance level of the mission's target, serving as the dependent variable in our analysis. The categories include:
  - `Primary target`: Highest priority targets critical to enemy's war-sustaining capabilities.
  - `Secondary target`: Important targets with significant, but not critical, impact.
  - `Target of opportunity`: Targets engaged opportunistically without prior planning.
  - `Target of last resort`: Targets chosen as a fallback when preferred options are not available.
- `target_industry`: The sector of infrastructure targeted, with the following possible values:
  - `Airfields`: Targeting enemy airstrips and aircraft facilities.
  - `Railway infrastructure`: Disrupting transportation lines and logistics.
  - `Synthetic oil refineries`: Striking at the heart of enemy fuel production.
  - `Unidentified targets`: Engaging targets that could not be specified.
  - `Urban areas`: Targeting urban centers with the potential to impact enemy morale or logistics.
- `country_mission`: The nation from which the mission originated, which can include:
  - `USA`: Missions flown by the United States.

– `Great Britain`: Missions flown by the United Kingdom.
– `Others`: Missions flown by other Allied nations.

- `bomb_tons`: The total weight of bombs dropped, reflecting the scale of the attack, with measures ranging from less than a ton to several hundred tons.
- `aircraft_attack`: The number of aircraft participating in the mission, a reflection of the committed resources and operational scale, ranging from single aircraft to entire squadrons.

Detailed information about these variables and the data structure is presented in Table 1, which outlines the first few records from the processed dataset.

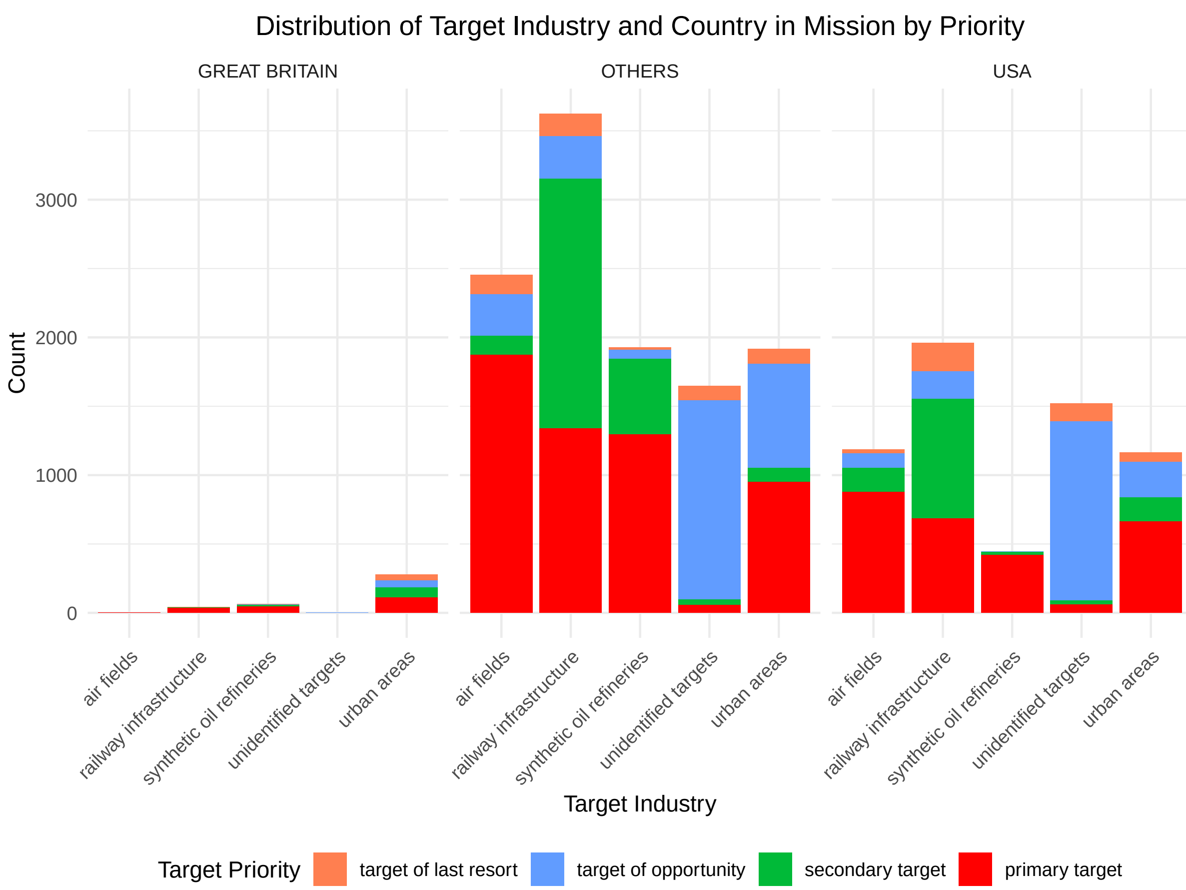


Figure 1: Distribution of target industry and country in mission by Priority

Figure 1 provides a extensive view of the processed data, showcasing the distribution of aerial bombardments across different industrial sectors targeted by the Allies during World War II, with a particular focus on those aimed at German industries. The stacked bar chart breaks down the count of missions by the leading country—Great Britain, USA, and others—and prioritizes the targets into four categories: primary target, secondary target, target of opportunity, and target of last resort.

From the visualization, we can discern that airfields and urban areas were frequently targeted, but with varying levels of priority depending on the mission's country of origin. The significant height of bars in categories such as synthetic oil refineries and railway infrastructure under the USA banner indicates a strategic emphasis on crippling Germany's logistic and production capabilities, aligning with the Allies' broader objectives of disrupting the war-sustaining resources of the Axis power.

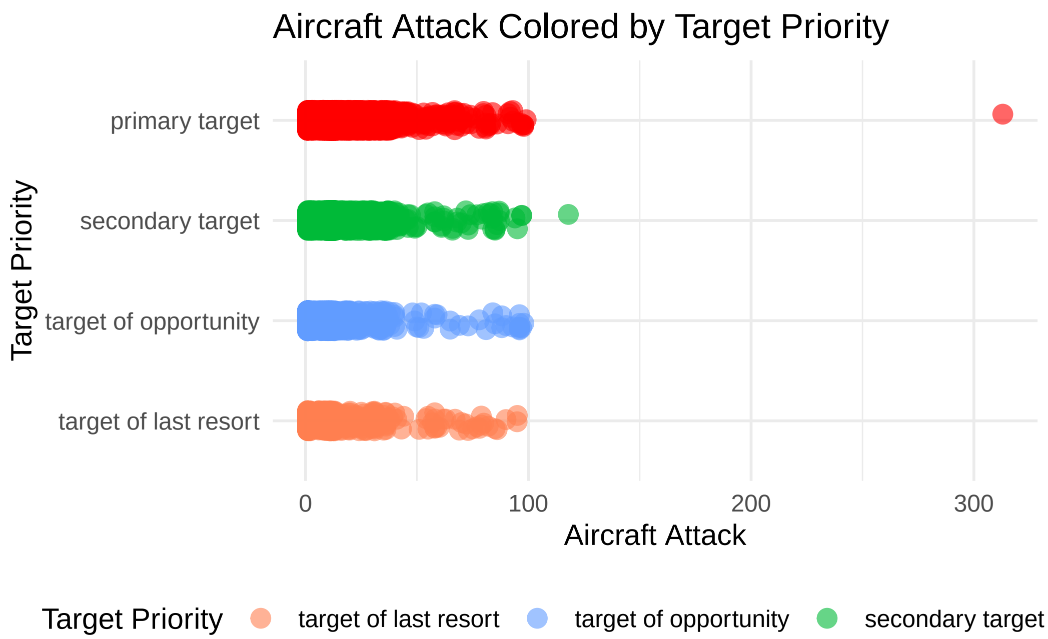


Figure 2: Distribution of aircrafts by Priority

Figure 2 depicts a distribution of aircraft attacks during World War II, colored by the priority of the target they were assigned to hit. The data points, each representing a specific mission, are spread across four categories of target priority: primary target, secondary target, target of opportunity, and target of last resort. The horizontal axis shows the number of aircraft involved in each attack, offering a visual correlation between the scale of the attack and the importance of the target. It's clear from the clustering of data points that primary targets often saw a higher concentration of aircraft, whereas targets of last resort were engaged with fewer aircraft, indicating a strategic allocation of aerial resources based on the priority of the target.

Figure 3 visualizes the weight of bombs dropped on various targets, categorized by the priority level of the mission. The volume of ordinance delivered is marked along the horizontal axis, indicating the intensity of each mission. Data points are color-coded to represent the different

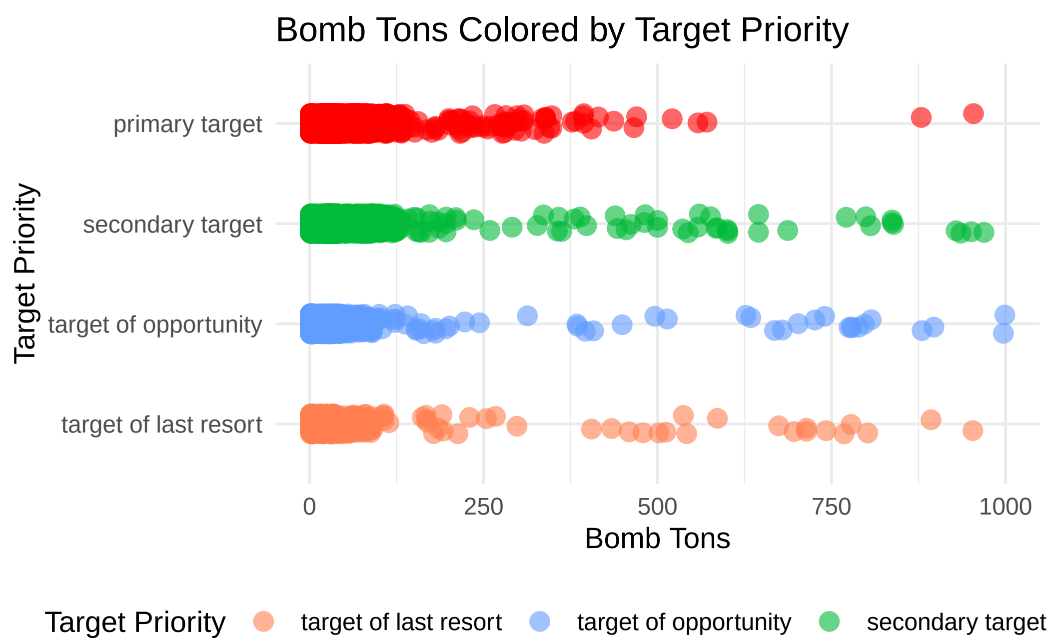


Figure 3: Distribution of bomb tons by Priority

priority levels: from targets of last resort, which saw the lightest bombardment, up to primary targets, which typically experienced the heaviest bombings. This spread suggests that higher-priority targets were often subjected to more intensive bombing campaigns, a strategic choice likely reflecting their importance to the war effort. The distribution illustrates the heavy investment in munitions for high-priority targets, reinforcing the critical nature of these missions.

2.3 Justification

The variables chosen for this study were carefully selected based on their strategic significance during World War II, as documented in the THOR dataset. Each variable offers insights into different aspects of the aerial bombing campaigns:

- `**Target Industry (TGT_INDUSTRY)**`: Identified based on the strategic value, targeting industries critical to the war effort, such as oil refineries and transportation networks.
- `**Country of Mission (COUNTRY_FLYING_MISSION)**`: Considered to analyze differences in target prioritization across the Allied nations.
- `**Total Tons of Bombs (TOTAL_TONS)**`: Used to gauge the intensity and importance assigned to each mission.
- `**Aircraft Attack (AC_ATTACKING)**`: Reflected the scale of the mission and was indicative of the level of commitment to the target.

Variables such as THEATER, TGT_TYPE, and MSN_TYPE were not included due to the presence of extensive missing values or the complexity of their classifications, which could complicate the analysis. Similarly, variables like TOTAL_LBS, TIME_OVER_TARGET, AC_LOST, and AC_DAMAGED were omitted to keep the model focused and interpretable.

To streamline the modeling process and enhance the clarity of the analysis, non-German targets were excluded from the original dataset of approximately 178,282 records. This data cleaning step resulted in a more focused dataset, comprising 64,948 entries, which facilitated a more efficient and targeted modeling approach. This pre-processing not only improved the manageability of the data but also honed the study's examination of the bombing strategies against Germany, offering a clear view of the Allied prioritization and tactical decisions.

3 Model

The goal of our modelling strategy is twofold. F This study employs a Bayesian ordered logistic regression model to analyze the priorities assigned to different target industries during aerial bombing missions in World War II.

The model integrates several predictors including the type of target industry (`tgt_industry`), the country executing the mission (`country_mission`), the tonnage of bombs dropped (`bomb_tons`), and the number of aircraft involved in the attack (`aircraft_attack`).

The model uses a logistic cumulative link function, draws on a posterior sample size of 4000, and is based on a total of 5000 observations, utilizing the rstanarm package to accommodate the ordinal nature of target priority levels.

Our model statistically infers the relationship between target priority and various strategic factors, providing a probabilistic assessment of their impacts.

Background details and diagnostics are included in Appendix A.

3.1 Model set-up

Let y_i be the ordered categorical variable representing the priority level of the target for the i-th bombing mission. The predictors in the model include:

- β_1: The coefficient for the `target_industry`, which indicates the type of industry targeted, such as oil refineries or transportation networks. It is a categorical variable with the following categories:
  - `Airfields`
  - `Railway infrastructure`
  - `Synthetic oil refineries`
  - `Unidentified targets`

  - Urban areas

- β_2: The coefficient for the `country_mission` variable, representing the country that carried out the mission. The possible values are `USA`, `Great Britain`, and `Others`.
- β_3: The coefficient for the `bomb_tons`, which is a continuous variable measuring the total tonnage of bombs dropped during the mission.
- β_4: The coefficient for the `aircraft_attack`, which is a count variable reflecting the number of aircraft participating in the mission.

Each coefficient β_j corresponds to the effect of the j-th predictor on the log-odds of the mission having a higher target priority.

- η_i: The linear predictor or log-odds for the i-th observation. It is a combination of the intercept and coefficients multiplied by the predictor variables.
- κ: The set of thresholds (or cutpoints) that define the boundaries between the ordered categories of the dependent variable. In the ordered logistic regression model, these thresholds determine the ranges of η_i values that correspond to each category of the target priority. The model estimates separate κ values for each boundary between the ordered categories.

$$
\begin{aligned}
y_i &\sim \text{OrderedLogistic}(\eta_i, \kappa) && (1)\\
\eta_i &= \beta_{\text{tgt\_industry}} \times \text{tgt\_industry}_i + \beta_{\text{country\_mission}} \times \text{country\_mission}_i && (2)\\
&\quad + \beta_{\text{bomb\_tons}} \times \text{bomb\_tons}_i + \beta_{\text{aircraft\_attack}} \times \text{aircraft\_attack}_i && (3)\\
\beta &\sim \text{Normal}(0, 10) \text{ (default non-informative prior)} && (4)\\
\kappa &\sim \text{Normal}(0, 5) \text{ (default prior for cutpoints)} && (5)
\end{aligned}
$$

3.2 Prior distributions

In the Bayesian ordered logistic regression model implemented using the `rstanarm` package, default priors are applied to the model parameters to ensure robust and reliable inference. These priors are designed to be weakly informative, balancing the need for regularization with the flexibility to adapt to the data:

- **Intercept Priors**: For the model's intercepts, a normal prior distribution is typically used with a mean of 0. This choice helps stabilize the location parameter without imposing too strong a belief about where it should be centered.

- **Coefficient Priors**: Coefficients in the model are also assigned normal prior distributions with a mean of 0. The standard deviation for these priors is often set around 2.5, a value chosen to prevent overly large effects unless strongly supported by the data, thereby adding a level of regularization.

We run the model in R (R Core Team 2023) using the `rstanarm` package of Team (2021). We use the default priors from `rstanarm`.These priors are intended to provide a default level of smoothing and regularization to the model, making it applicable to a wide range of datasets while still being flexible enough to capture essential data-driven insights.

4 Results

4.1 Model Justification

The analysis revealed distinct patterns in target prioritization, highlighting the role of industry type and the mission's originating country in determining target selection, while the impact of tonnage and aircraft numbers was relatively subdued. Table 2.

Table 2: The model's coefficient summary

Parameter	Mean	SD	10%	50%	90%
Railway infrastructure	-1.4	0.1	-1.6	-1.4	-1.3
Synthetic oil refineries	-0.1	0.1	-0.2	-0.1	0.1
Unidentified targets	-3.4	0.1	-3.5	-3.4	-3.2
Urban areas	-1.4	0.1	-1.6	-1.4	-1.3
Other countries	0.8	0.3	0.4	0.8	1.2
USA	0.9	0.3	0.5	0.9	1.3
Bomb	0.0	0.0	0.0	0.0	0.0
Aircraft	0.1	0.0	0.1	0.1	0.1
Intercept 1 (last resort|opportunity)	-3.7	0.4	-4.1	-3.7	-3.2
Intercept 2 (opportunity|secondary)	-1.0	0.3	-1.4	-1.0	-0.5
Intercept 3 (secondary|primary)	0.4	0.3	0.0	0.4	0.9

As detailed in Table 2, the coefficient summary quantitatively reflects the strategic emphases placed on various targets during World War II aerial bombings. For instance, the estimated coefficient for `Unidentified targets` is notably negative (Mean = -3.4), suggesting a strategic de-emphasis on these targets in comparison to other types.

Conversely, `USA`-led missions are associated with a positive coefficient (Mean = 0.9), indicating a higher probability of such missions targeting high-priority objectives. This aligns with the historical context of the USA's significant role in the strategic bombing campaign. The coefficient for `Other countries` is marginally lower (Mean = 0.8), subtly reflecting the diverse strategic approaches among the Allied forces.

The model's intercept terms serve to baseline the log-odds of the priority classifications, ranging from `Intercept 1 (last resort|opportunity)` with a mean of -3.7 to `Intercept 3 (secondary|primary)` with a mean of 0.4. These intercepts delineate the inherent ordering of target priority levels within the model, with higher intercepts corresponding to a higher likelihood of a target being deemed of primary importance.

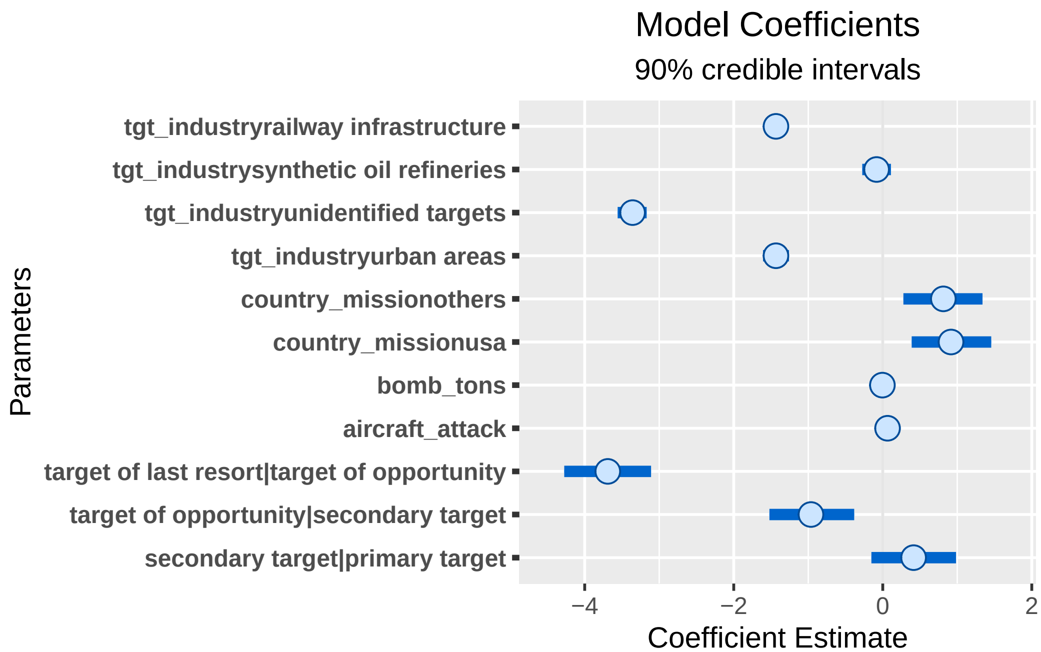


Figure 4: The 90% credible intervals for all model coefficients

The Bayesian ordered logistic regression model's estimates are visualized in Figure 4. Each point in the plot represents the posterior mean effect size of the predictor variables on the target priority ranking, while the lines indicate the 95% credible intervals. The estimates reveal several key insights into the Allied bombing strategy:

- The variable `tgt_industrySynthetic Oil Refineries` exhibits the largest positive effect, suggesting that synthetic oil refineries were assigned the highest priority for bombing missions, in line with the strategic objective to disrupt the German war effort's fuel supply.
- In contrast, the effects associated with `tgt_industryRailway Infrastructure` and `tgt_industryUrban Areas` are closer to zero, indicating a lower priority relative to other types of targets. This aligns with a strategy that placed less emphasis on disrupting transportation and civilian structures.

- Regarding the country of mission execution, `country_missionusa` shows a significant positive effect, reflecting the United States' prominent role in the strategic bombing campaign against Germany.
- The effects of `bomb_tons` and `aircraft_attack` are positive, suggesting that missions with greater bomb tonnage and more aircraft tend to have a higher target priority, possibly reflecting the allocation of resources to strategically important missions.

These findings underscore the strategic considerations that guided the Allies' targeting decisions during World War II, highlighting the emphasis on industrial targets critical to the German war machine.

5 Discussion

This paper has initiated an inquiry into the labyrinth of decision-making within the Allied bombing strategy of World War II, applying a Bayesian ordered logistic regression model to discern the prioritization of targets. By selectively incorporating a subset of variables available in a rich dataset, the study uncovers patterns that delineate strategic imperatives from the vast array of aerial operations data.

5.1 Extensive Understanding of Target Selection

In this study, key variables such as target industry, country flying mission, bomb tons, and aircraft attacking were carefully selected based on their demonstrated strong correlation with target priority. This choice was guided by historical and strategic relevance, ensuring that the model reflects real-world considerations. The quantitative analysis substantiates the critical impact of these variables on strategic military decisions, revealing how different target industries were prioritized based on their role in the war effort. For instance, industries essential to the enemy's war logistics, such as oil refineries, were often prioritized to cripple the enemy's ability to sustain military operations. This not only aligns with historical records but also provides a granular insight into the decision-making processes, illustrating the careful balancing act between achieving strategic objectives and managing resources effectively during the war.

5.2 Strategic Implications of Variable Selection

The selection of variables for this analysis was carefully orchestrated to capture the most salient elements of the strategic bombing campaigns, aligning with historical accounts and operational records. In particular, the emphasis on target industry and mission origins provides novel insights into the resource allocation and strategic planning of the Allied forces. By scrutinizing these variables, this study sheds light on the underlying strategic motives, revealing a pattern of prioritization that favored crippling the enemy's production capacity over other targets.

While the exclusion of certain variables, such as THEATER and TGT_TYPE, was a deliberate decision to ensure data integrity, it also opens up avenues for future research. Subsequent studies could explore how different theaters of war might have diverged in their strategic approach to bombing campaigns, or how specific target types contributed to the overall war effort. These aspects, while beyond the scope of the current analysis, suggest a rich tapestry of strategic considerations that could be further unraveled.

5.3 Weaknesses and Future Research Directions

The scope of this study, while extensive, is constrained by the variables selected and the overall dataset size. The omission of variables such as THEATER and TGT_TYPE, although necessary to maintain data quality, limits the depth of strategic analysis possible. Future studies could benefit from incorporating these and other variables, such as TIME_OVER_TARGET and AC_LOST, to provide a fuller picture of the strategic complexities of World War II aerial campaigns.

Moreover, while the model provides a solid foundation, its predictive accuracy could be improved through more advanced sampling methods or by integrating additional datasets that offer broader temporal and geographic coverage. For instance, expanding the dataset to include more varied mission types and broader operational theaters could offer new insights into the adaptive strategies employed by the Allies over the course of the war.

Looking ahead, there is significant potential to examine the effects of bombing strategies on civilian populations by integrating geographical data. This could help in understanding the broader humanitarian impacts of strategic bombing. Additionally, exploring the relationship between flight altitude and mission success could yield valuable insights into the tactical decisions made during these missions. This line of inquiry could benefit from the application of advanced statistical techniques or machine learning models to uncover patterns not readily visible through traditional methods.

Each of these directions not only promises to expand the current understanding of wartime strategies but also contributes to the broader discourse on the ethical and strategic implications of aerial warfare both historically and in contemporary contexts.

5.4 Envisioning the Future of Historical Military Analysis

Prospects for future research based on this dataset are immense. One could investigate whether there were unwritten standards during World War II for minimizing civilian harm, using the geographical coordinates available in the data to map bombing patterns against civilian population centers. Such analysis could offer new insights into the ethical boundaries of strategic warfare.

Furthermore, understanding the correlation between flying altitude and success rates could clarify tactical nuances that determined mission outcomes. The wealth of information contained in the dataset paves the way for such multifaceted investigations.

5.5 The Value of Strategic Insights

This research underscores the analytical value of prioritization patterns, such as the focus on oil refineries, which could be extrapolated to anticipate enemy targets and inform defensive strategies. The prominence of American-led missions in prioritization further reflects the geopolitical dynamics of World War II, providing a statistical testament to the United States' influential role during the conflict.

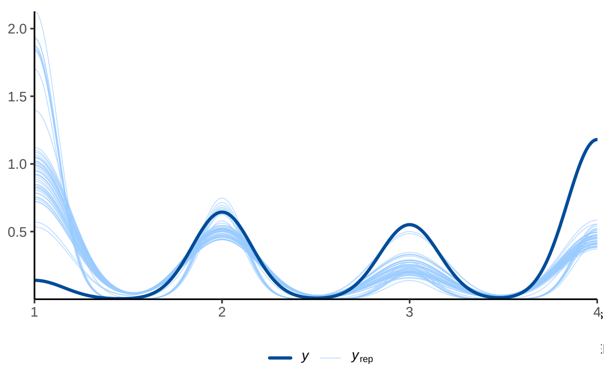
Appendix

A Model details

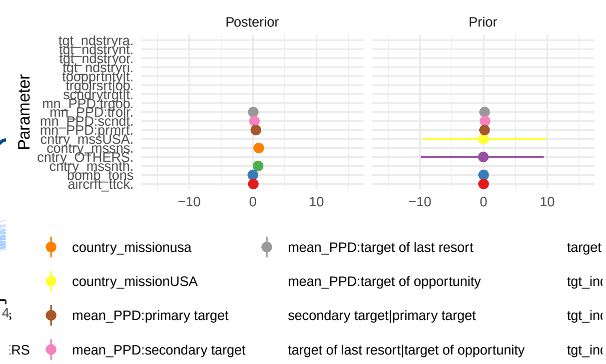
A.1 Posterior predictive check

In Figure 5a we implement a posterior predictive check. This shows a posterior prediction check, which compares the observed data (denoted by y) against the replicated data generated by the model (denoted by y__rep). The overlaid lines represent multiple posterior predictive distributions, providing a visual assessment of how well the model predictions align with the actual observed data across the range of predicted values. The close alignment between the curves indicates a good model fit, as the simulated data appears to capture the variability and central tendency of the observed data.

In Figure 5b we compare the posterior with the prior. This shows a strong influence of the data on the parameter estimates. This plot is essential for understanding how the evidence in the data alters prior assumptions to shape the final inference drawn from the Bayesian model.



(a) Posterior prediction check



(b) Comparing the posterior with the prior

Figure 5: Examining how the model fits, and is affected by, the data

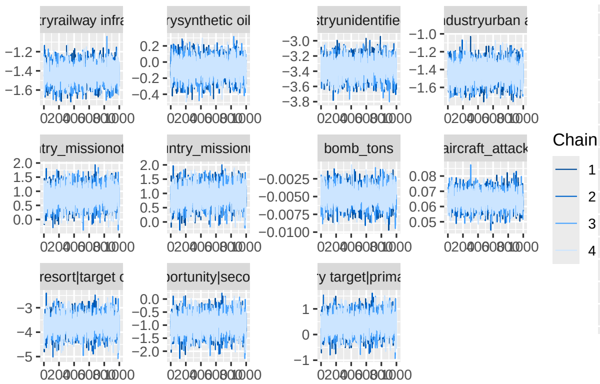
A.2 Diagnostics

Figure 6a is a trace plot. It shows how the sampled values of each parameter evolve over time. Ideally, a well-mixed chain will resemble a 'hairy caterpillar', indicating that the sampling has explored the posterior distribution extensively and has likely achieved convergence.

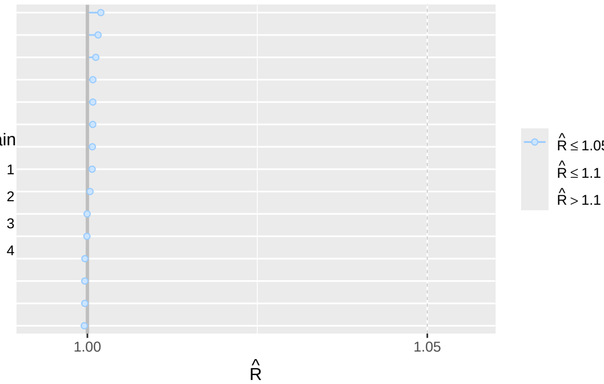
This suggests that the chains for each parameter have likely converged to their respective posterior distributions. This is a critical aspect of ensuring the reliability of the Bayesian estimates for these parameters. The plot, therefore, supports the credibility of the posterior estimates derived from the model, provided that other diagnostic checks also affirm convergence.

Figure 6b is a Rhat plot. It shows that the chains have converged to a common distribution, which is an indication of good mixing and reliable posterior estimates.

This suggests convergence has likely been achieved, and the posterior distributions can be considered trustworthy for inference.



(a) Trace plot



(b) Rhat plot

Figure 6: Checking the convergence of the MCMC algorithm

Figure 7 is a graph shows the posterior distributions of the parameters from a Bayesian ordered logistic regression model. Each horizontal line represents the 50% credible interval, centered around the median of the posterior distribution for a given parameter, with the ends of the lines marking the 25th and 75th percentiles. The length of each line indicates the degree of uncertainty associated with the estimate of that parameter.

The parameters include various target industries, countries of mission, and other factors like bomb tonnage and aircraft attacks. Notably, parameters such as tgt_industryunidentified targets exhibit a more negative median value, which might suggest a lower priority in the bombing strategy compared to other target types.

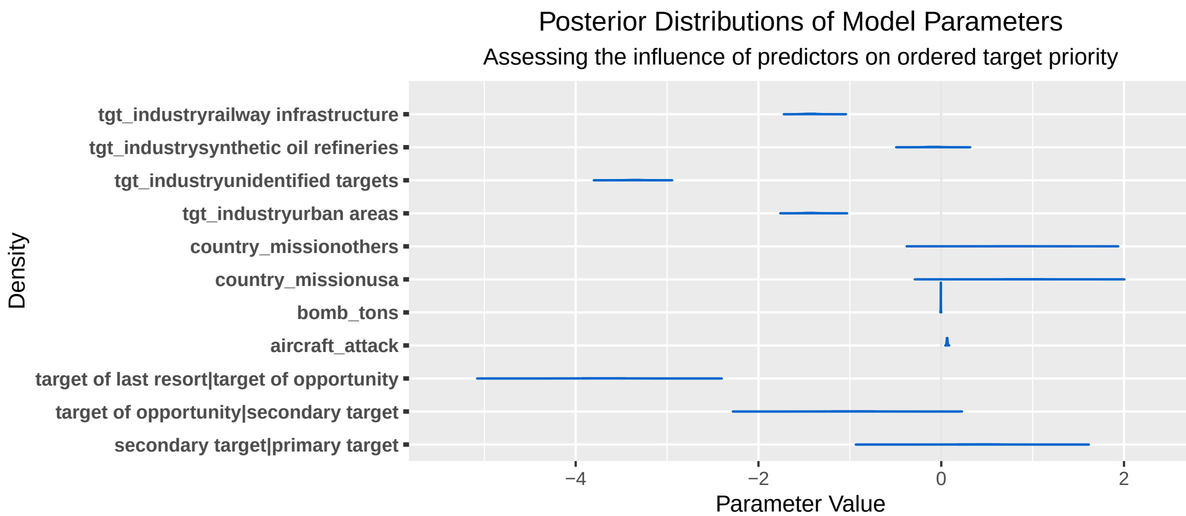


Figure 7: The posterior distributions for all the parameters

References

Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. https://tellingstorieswithdata.com.

Apache Arrow. 2021. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

Biddle, Stephen. 2002. *Military Power: Explaining Victory and Defeat in Modern Battle*. Princeton University Press.

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gelman, Gabriel, Jonah Gabry, et al. 2021. *Bayesplot: Plotting for Bayesian Models*. https://mc-stan.org/bayesplot.

Grey, Colin. 1999. *A War of a Different Kind: Strategic Bombing in World War II*. Military History Press.

Morrow, James D. 2004. "The Strategic Setting of Choices: Signaling, Commitment, and Negotiation in International Politics." *Strategic Studies Quarterly*.

Overy, Richard. 2013. *The Bombers and the Bombed: Allied Air War over Europe, 1940-1945*. Viking.

Pape, Robert A. 1996. *Bombing to Win: Air Power and Coercion in War*. Cornell University Press.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Team, Stan Development. 2021. *Rstanarm: Bayesian Applied Regression Modeling via Stan*. https://mc-stan.org/rstanarm.

Wickham, Hadley. 2021. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of

Graphics. https://ggplot2.tidyverse.org.

Wickham, Hadley et al. 2021. *Tidyverse: Easily Install and Load the 'Tidyverse'*. https://tidyverse.tidyverse.org.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org.

Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.