Procurements of the Canadian Government, Runaway Military Contracts*

Robert Ford

December 2, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

In recent years, Canada's federal spending landscape has undergone significant transformations, particularly with a marked increase in military expenditures. The government has announced substantial new military spending packages, [REF]

Federal procurements involve the acquisition of goods and services by government departments, Buyers, from companies, Suppliers. These procurements are essential for a wide range of government functions from; acquiring defense equipment, building infrastructure, procuring healthcare supplies, technological services and engineering consultants for example. This paper will analyse federal procurement data to dissect government procurements over the last four years.

The primary questions this paper aims to address are: What is the government buying? Who are the primary suppliers? How are spending patterns shifting across different departments and over time?

I collected and processed detailed data on over 301,000 contracts awarded since January 2020. The dataset includes variables such as contract amounts, durations, award dates, Buyers, and Suppliers. We derived additional metrics to facilitate our analysis, such as Baseline Days (the number of days since January 1, 2020) and Duration Days (the length of each contract), and standardized supplier names to ensure consistency.

Some significant findings presented in this paper include: military procurement spending has surged in the past few years. Where the Department of National Defence now dominates federal procurement spending, with expenditures exceeding $56 billion since 2020. This increase in

*Code and data are available at: https://github.com/Ford-Robert

spending is in part thanks to large contracts, including a $11.2 billion contract to SkyAlyne for fighter pilot training—the largest in our dataset—and multi-billion-dollar contracts for fighter jets and naval vessels.

While other government departments have experienced a notable decrease in procurement spending. For instance, the Public Health Agency of Canada's spending, which was substantial in 2020 due to the COVID-19 pandemic, has declined significantly in subsequent years.

These findings are important for several reasons. They provide critical insights into how government spending priorities are shifting, which has implications for economic policy, industry stakeholders, and public transparency. Understanding the concentration of spending in large military contracts can inform discussions on national defense strategies, budget allocations, and the impact on domestic industries.

The paper is structured as follows: [Section 2] details the data collection and processing methodologies, including the sources of procurement data and the steps taken to clean and standardize the dataset. Then I present a series of graphs to demonstrate government procurements from various angles. [Section 3] describes the linear regression models used to assess trends in contract amounts over time and contract duration. [Section 4] presents the results of the analysis, highlighting key trends in military and non-military procurement spending. [Section 5] discusses the limitations of the study, including data constraints and methodological considerations, and suggests areas for future research. Finally, Section 6 concludes with a summary of the findings and their implications for policymakers, industry stakeholders, and future studies.

By providing a comprehensive analysis of federal procurement spending, this paper aims to fill the existing gap in understanding how government consumption of private sector goods and services is changing over time. The insights gained from this study can inform policy decisions, contribute to academic discourse, and enhance transparency regarding government spending practices.

TODO Cite R and all other packages

TODO Estimand?

2 Data

The data was collected using R and other packages [REF]. A web scraping script was used to collect all Federal Award Data from the IJF [REF]. The variables this paper focuses collected from the IJF are; Contract, Buyer, Supplier, Amount, Award Date, Start Date, and End Date.

Contract: A few words detailing a brief description of what the contract signed was for

Buyer: The name of government department who awarded the contract

Supplier: The name of the company that fulfilled the contract by agreeing to provide the good or service to the Buyer

Amount: The value of the contract in Canadian Dollars

Award Date: The day the Buyer awarded the contract to the supplier

Start Date: When the contract is due to start

End Date: When the contract is due to end

I derived three additional variables using the IJF variables:

Baseline Days: Number of days since 1st January 2020

Duration Days: End Date subtracted from the Start Date to calculate the duration of the contract

Processed Supplier: The name of suppliers regularized to allow for aggregation and comparison

Baseline Days and Duration Days are used as predictors in the linear model, as Dates are not valid entries into a linear model.

Processed Supplier is required as many companies have slightly different spellings, such as ["General Dynamics", "GENERAL DYNAMICS", "General Dynamics Mission Systems", "GENERAL DYNAMICS MISSION SYSTEMS INC"]. See Appendix TODO for how I processed and regularized the suppliers.

Here is the head of the data that was used in this analysis for a sense of what these variables represent. TODO Fix table

Table 1: Sample of the Data

region	buyer	supplier	amount	awarc
Federal	Canadian Centre for Occupational Health and Safety	IPSS Inc	39958	2024-
Federal	Canadian Centre for Occupational Health and Safety	Articulate Global LLC	22785	2024-
Federal	Canadian Centre for Occupational Health and Safety	Colony Labs, dba Scribe	12672	2024-
Federal	Canadian Centre for Occupational Health and Safety	Sheraton Suites Calgary Eau Claire	100000	2024-
Federal	Canadian Centre for Occupational Health and Safety	ROSS & McBRIDE LLP	20000	2024-

Figure 1: template

The data was then cleaned. First I began by setting each variable to the correct data type. Then I removed any data point where Amount was negative or zero, because I am not sure what a negative amount means in this context and contracts that are not worth anything are not worth looking at. There were a number of contracts which started and ended on the same day, to simplify later calculation I simply added a day to their End Date, figuring

that same day contracts are equivalent to one day contracts. Further data cleaning details are available in Appendix TODO. After cleaning I was left with 301,094 contracts spanning TODO to TODO.

2.1 Measurement

The data is recorded and published through various government agencies, who use different accounting standards and procedures. The three sources IFJ used to collect the federal procurement data are; Canada Buys, Buy and Sell, and Proactive Disclosures (which makes up the vast majority of the procurement data). However, because the data used in this paper is pulled from the IFJ's Federal Awards page this data comes exclusively from the Proactive Disclosures. The Treasury Board of Canada Secretariat is responsible for publishing these Proactive Disclosures under the Access to Information Act [REF].

2.2 Analysis

This is a break down of total amount awarded by the top 5 biggest spending Buyers. The Other category consists the of the total awards given by the remaining 88 smallest Government Buyers. It is clear that National Defence dwarfs the procurement budget, spending over /$56 Billion on procurements since 2020.

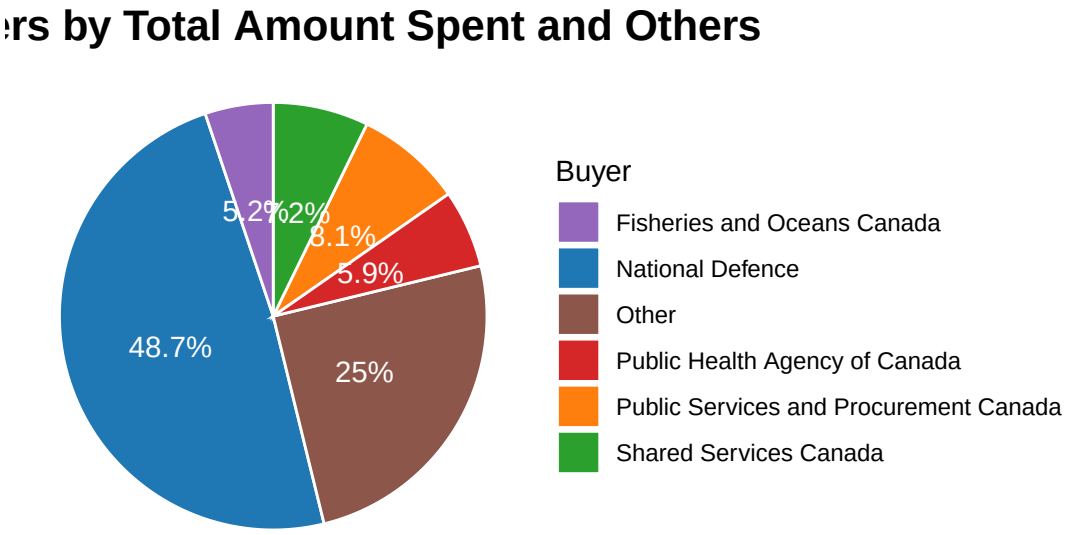

Figure 2: Top 5 Buyers by Total Amount Spent and Others (Pie Chart)

FIG show how National Defence spending has taken off in recent years, where is 2021 it had not even crested /$5 Billion, but in the last 3 years spending has exploded to /$15 Billion a year.

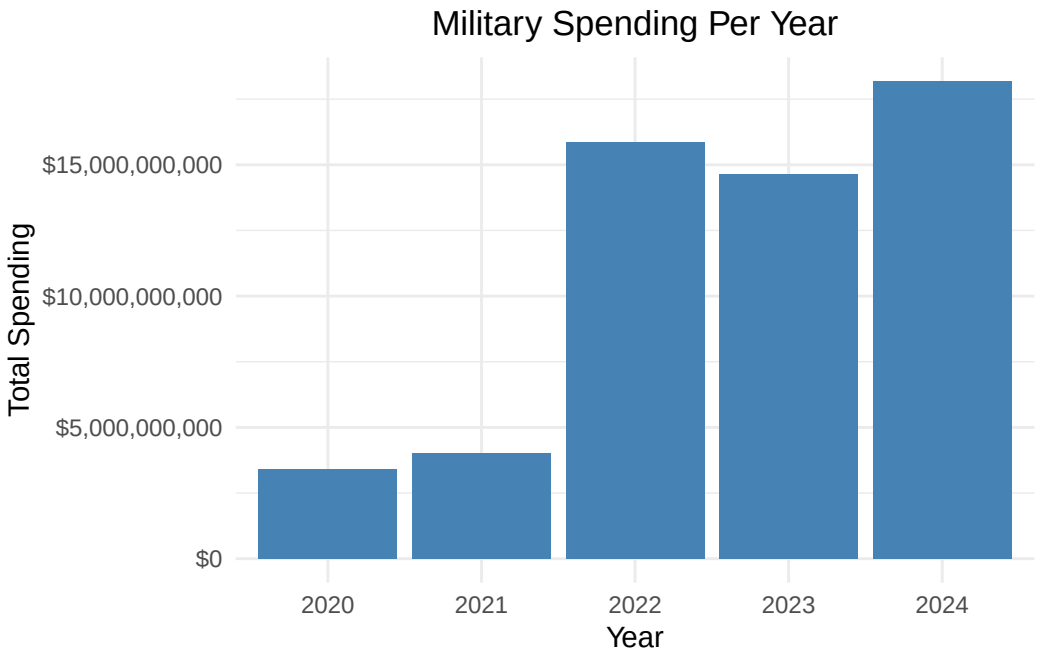

Figure 3: template

Public Services and Procurements Canada is the clearinghouse for the entire federal government. This means they are responsible for managing the procurement process for buyers across the government [REF]. It is important to note that the PSPC sometimes does not publish small contracts, less than /$40,000. PSPC spending over time shows while military spending increases the rest of the government is taking a hit.

This FIG shows the top five Buyers spending per year since 2020. In the first 2 years military spending held steady and did not dominate the budget by any means. But since then Defence spending exploded not just nominally but compared to the rest of the budget as well. Notably the Public Health Agency of Canada's spending while strong in 2020, decreased to nothing over the following four years. The spending in 2020 is linked to purchases of medical equipment, so it is possible they have not needed to invest in new equipment since then. But this illustrates a boarder trend among all departments except National Defence, where their spending has steadily declined in the last 4 years.

To find out the firms that are benefiting the most through this procurement process we can look to FIG. Unsurprisingly, the biggest suppliers are primarily are awarded contracts by the military. Interestingly the top supplier, Skyalyne, was only awarded one huge 11.2 Billion contract for fighter pilot training services. This is the largest contract in the dataset. Then considering the largest contracts awarded in the last 3 years to; the F-35 program (Fighter Planes, 4.5 Billion), General Atomics (Airplanes, 1.7 Billion), CAE (Aviation Engineering and Training, 4.4 Billion), Airbus (Airplanes, 3.7 Billion), for a grand total of 25.5 Billion. These large single contracts devoted to advanced fighter jets and training pilots to fly them is part

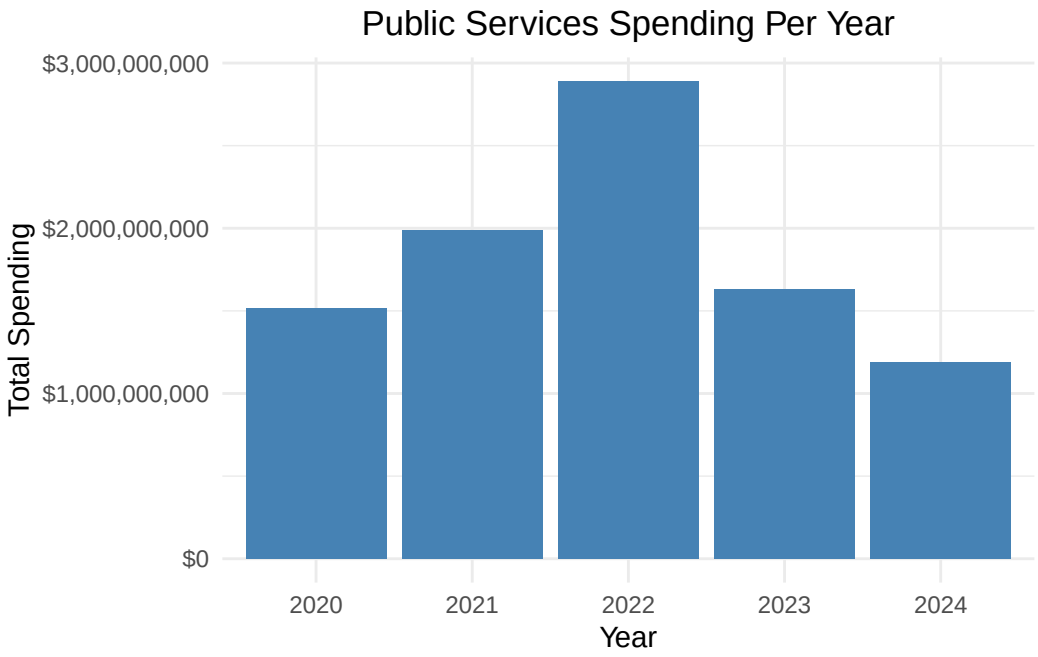

Figure 4: template

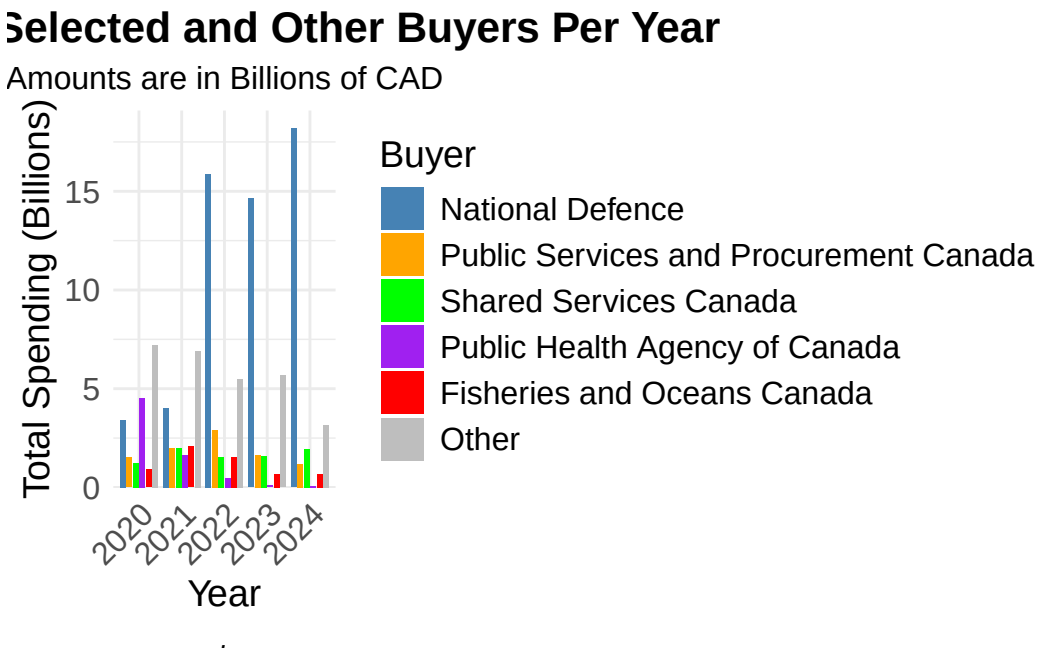

Figure 5: template

of the reason that military spending has exploded in the last 3 years. The military also paid Vancouver Shipyards 3.4 Billion in 2024 for Ships and Boat parts.

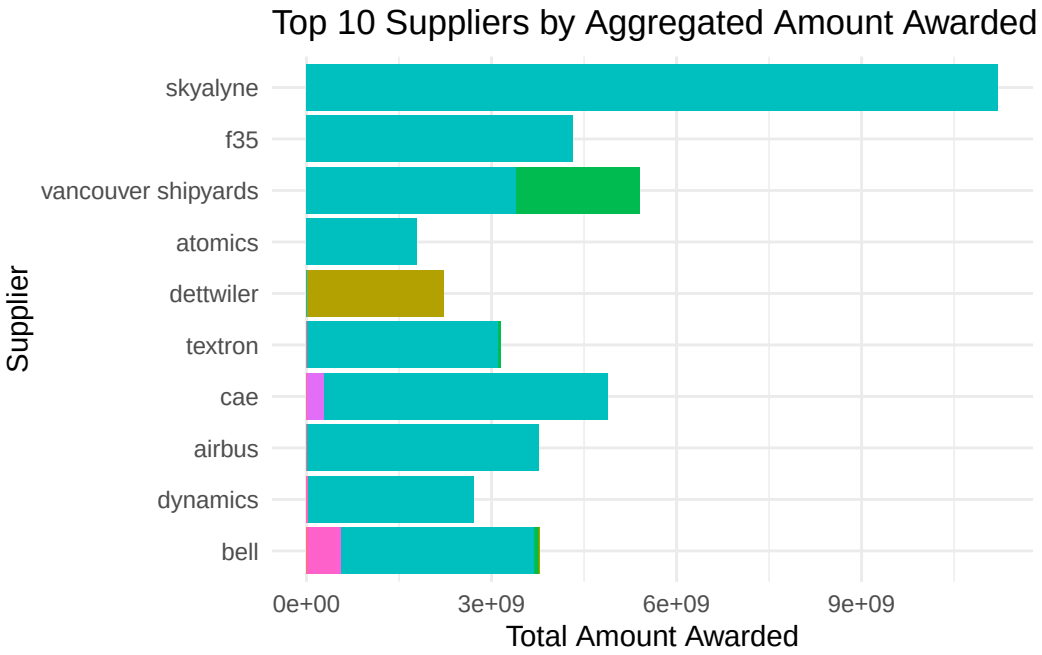

Figure 6: template

Dettwiler is an engineering firm that provides the Canadian Space Agency with engineering and scientific consulting. Surprisingly 86.3% of the Space Agency's procurement budget goes to Dettwiler.

2.3 Distribution

This graph represents 81.75% of all contracts, indicating that the vast majority of contracts awarded by the government are quite small. Furthermore, it seems that the number of contracts follows a Pareto distribution. There are 102 contracts that are worth $1, this maybe some legal formality.

```
[1] "Percentage of contracts <= 100,000 CAD: 81.75%"
```

Though most contracts are small contracts (less than $100,000), the sum of these 246,000 contracts is only worth a measly 7.34 Billion. The Skyalyne contract itself outweighs all the small contracts.

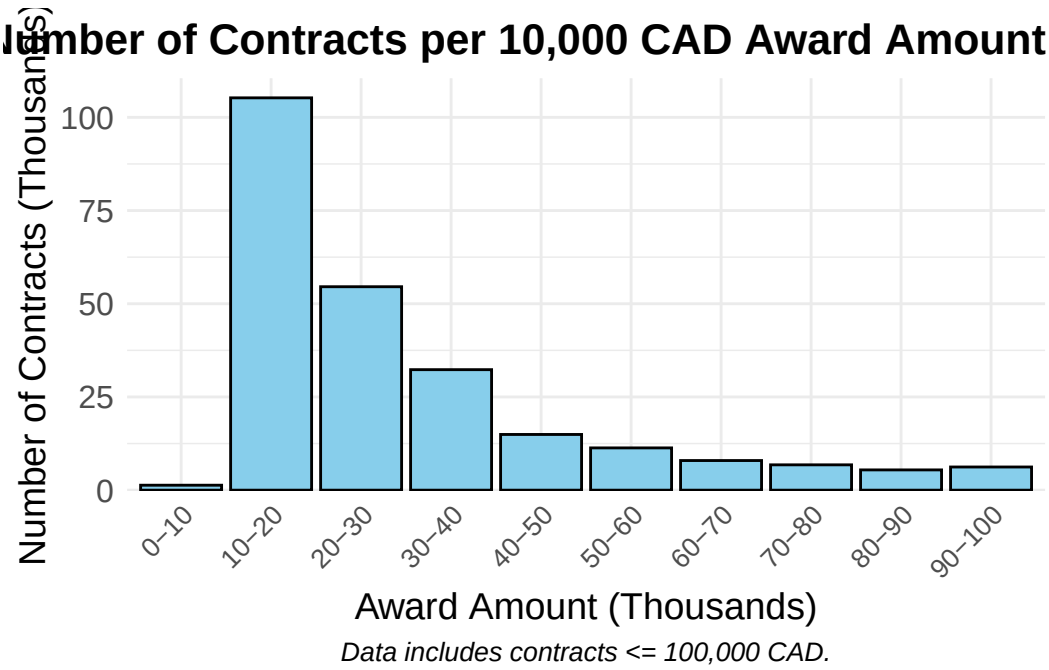

Figure 7: template

```
# A tibble: 2 x 3
  contract_size            total_amount_spent total_amount_billion
  <chr>                                 <dbl>                 <dbl>
1 Big Contracts (> 100k)         108296409037                108.
2 Small Contracts (  100k)         7344921521                  7.34
```

3 Model

This paper will use a simple linear regression to better understand the how contract award amounts are affected over time compared to their lengths. To do this the response variable will be Amount, and the two dependent variables will be the number of day since January 1st 2020 (start of the data set) and the length of the contract. The coefficients of these variables will indicate the relative importance of when the contract was issues compared to the length of contracts on determining the size of contracts. I will create two models, one to measure the coefficients for all the contracts excluding military contracts and a model for solely military contracts. This way we can compare to see if military contracts are getting larger because they are relatively longer or because recently they have purchased more expensive equipment.

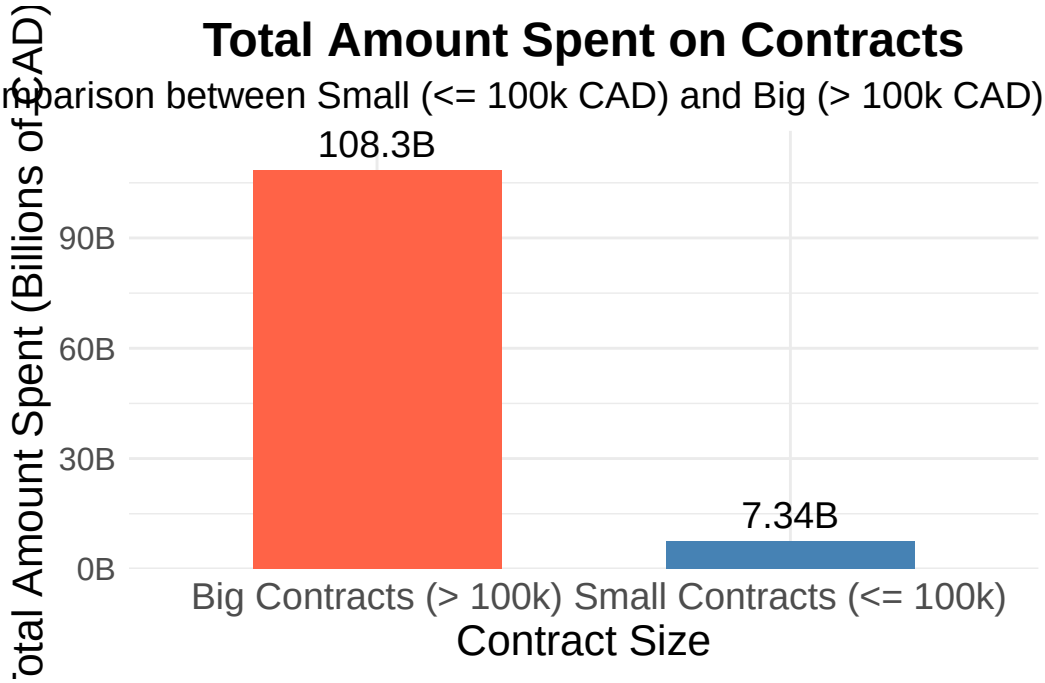

Figure 8: template

3.1 Model set-up

All model construction and diagnostics were conduction using R [REF], and the following packages:

Tidyverse: For data manipulation and visualization [REF]

Dplyr: [REF]

Car: [REF]

Caret:[REF]

Model Formula:

TODO

Variables:

Amount:

Baseline Days:

Duration Days:

3.2 Limitations

One significant limitation is the assumption of a linear relationship between the predictors and the contract amount. In reality, the relationship may be more complex or nonlinear, especially over an extended period where economic factors and policy changes could introduce fluctuations not captured by a simple linear trend.

Another limitation is the potential omission of relevant variables that influence contract amounts. Factors such as inflation, changes in government procurement policies, economic cycles, or industry-specific trends can significantly affect contract values. The exclusion of these variables may lead to omitted variable bias, where the estimated effects of the included predictors are distorted because they are capturing the influence of missing factors.

Additionally, the presence of outliers or extreme values, particularly in financial data like contract amounts, can disproportionately influence the model estimates and lead to misleading conclusions.

This model treats all contracts as a homogeneous group. But considering that contracts both for National Defence and general government awards vary widely in scope and cost that would suggest a linear model is not appropriate. A more segmented model may be appropriate. This may be difficult as dividing the data in to groups based on contract size would not be sufficient as there is still a wide variety of contracts that are of similar size. Possibly using the Contract descriptions to divide the data, but this would require more advanced Natural Language processing techniques and in many cases the descriptions are hardly descriptive.

Moreover, the model assumes that the effect of each predictor is constant over time. If the influence of contract duration or time on contract amounts changes due to evolving market conditions or government strategies, the model's fixed coefficients would not reflect these dynamic relationships.

3.3 Diagnostics

TODO

4 Results

The model general contracts indicates that the contract amount tends to decrease over time. Because Baseline days is slightly negative. This value suggests that every day that passes on average the value of non-military contracts is going down by $61.62 a day. This decline could reflect shifts in government spending priorities away from non-military departments. Additionally, as expect Duration Days is positive of the contract amount . This means that longer contracts are more expensive for the government.

In contrast, the model for military contracts reveals that the contract amount increases over time. The positive coefficient for Baseline days suggests an upward trend in military contract values since 2020. In this case the model suggests that the cost of military contracts is increasing by $852.50 every day. This is trend is due to the huge increases in military procurement we observed in FIG.

The impact of contract duration on the contract amount is substantially larger in military contracts compared to non-military ones. Where military contracts cost $38,146 more per day of the length of the contract. This is again due largely by the 11 Billon Skyalyne contract, and the other large long contracts for training and fighter jets.

Table 2: Model Results Summary

Model	Term	Estimate	Std. Error	t-Statistic	p-Value
Excluding Military Contracts	Intercept	-39457.6159	24751.90295	-1.594124	0.1109095
Excluding Military Contracts	Baseline Days	-61.6224	23.80006	-2.589170	0.0096214
Excluding Military Contracts	Duration Days	1341.7462	32.82717	40.873040	0.0000000
Military Contracts Only	Intercept	-5485218.7975	402365.90658	-13.632414	0.0000000
Military Contracts Only	Baseline Days	852.4927	393.51927	2.166330	0.0302892
Military Contracts Only	Duration Days	38145.9891	677.11532	56.336030	0.0000000

Figure 9: Model Results Summary

5 Discussion

5.1 First discussion point

This study conducted an in-depth analysis of Canadian federal procurement data, with a focus on military contracts. The data was collected and processed to ensure accuracy and relevance. Using web scraping techniques, we gathered information from the Institute for Government's (IJF) Federal Award Data. The primary variables included contract descriptions, buyer and supplier names, contract amounts, and key dates such as award, start, and end dates.

To facilitate the analysis, we derived additional variables: Baseline Days, representing the number of days since January 1st, 2020, and Duration Days, calculated by subtracting the start date from the end date of each contract. These transformations were necessary because date variables are not directly suitable for linear modeling. We also standardized supplier names through a Processed Supplier variable to ensure consistency, given the variations in how supplier names were recorded.

Data cleaning was a critical step. Contracts with negative or zero amounts were excluded due to their ambiguous nature. Contracts that started and ended on the same day were adjusted by extending the end date by one day to simplify duration calculations. After these refinements, the dataset comprised 301,094 contracts spanning from [TODO: start date] to [TODO: end date].

For the modeling phase, we employed a simple linear regression to explore how contract amounts are influenced by time and contract duration. Two separate models were constructed: one encompassing all contracts excluding military contracts, and another focusing solely on military contracts. This bifurcation allowed us to compare trends between military and non-military spending.

5.2 Second discussion point

The analysis revealed significant trends in Canada's federal procurement spending:

Surge in Military Spending: Military spending has dramatically increased over the past couple of years. National Defence emerged as the dominant spender, allocating over $56 billion since 2020. While military spending was under $5 billion in 2021, it escalated to approximately $15 billion annually in the following years. This spike is attributed to substantial contracts for advanced military equipment and services, such as the $11.2 billion contract awarded to SkyAlyne for fighter pilot training—the largest in the dataset. Other notable contracts include multi-billion-dollar deals with Lockheed Martin, General Atomics, CAE, and Airbus for fighter planes and aviation services.

Decline in Other Departments' Spending: In contrast, other government departments have experienced significant reductions in procurement spending. For instance, the Public Health Agency of Canada's spending, which was robust in 2020 due to pandemic-related expenditures, diminished to negligible levels in the subsequent years.

Prevalence of Small Contracts: The majority of government contracts are relatively small. Approximately 81.75% of all contracts are valued under $100,000, with the most common contract size ranging between $10,000 and $20,000. This pattern indicates a governmental preference for smaller contracts, possibly for regulatory reasons. There may be a rule that if a contract is below 20000 then it requires less scrutiny to be passed, in effect encouraging department to award more smaller contracts. Though it does not seem to be out of control as only TODO% of all spending in last 4 years is concentrated in contracts less than 100000.

5.3 Third discussion point

5.4 Weaknesses

While the study provides valuable insights, several limitations must be acknowledged:

Supplier Name Analysis: The methodology for standardizing supplier names, requires more rigorous techniques. The method used ensured that the top 10 suppliers are accurately aggregated. However, some contracts that these suppliers may have been missed. But beyond the top 10, the names of suppliers not well parsed out, with problems like variations of "incorporated" and that sort of thing, dominating the rest of the suppliers list.

Model Limitations: The linear regression model demonstrated weaknesses, primarily due to its simplicity and the limitations of the dataset. The assumption of a linear relationship between the predictors (Baseline Days and Duration Days) and contract amounts may not capture the complexities of procurement spending, which can be influenced by many things such as; policy changes, economic conditions, and sector-specific dynamics.

Methodological Constraints: The model did not account for external variables such as inflation, exchange rates, or changes in government procurement policies, which could significantly influence contract amounts. Additionally, treating all contracts as a homogeneous group overlooks the diversity in contract types and scopes.

5.5 Next Steps

To address these weaknesses and build upon the findings, the following steps are recommended:

Enhance Data Analysis Techniques: Supplier Name Processing: Implement advanced text processing techniques, such as fuzzy matching, to accurately standardize supplier names and improve data reliability. ADD MORE

Contract Description Analysis: Utilize text mining and content analysis methods on contract descriptions to uncover patterns in procurement objectives and priorities. ADD MORE

Expand Data Collection: Comprehensive Spending Data: Include other forms of government spending, such as in-house expenditures on salaries and operational costs, as this would contextualize the procurement budget in the larger scheme of federal spending overall. How much of the total federal budget goes into procurements? Are procurements more cost effective than in house spending?

Investigate Specific Trends: Public Health Agency Spending: Examine the factors contributing to the decline in the Public Health Agency of Canada's procurement spending.

Proliferation of Small Contracts: Analyze the rationale behind the high frequency of small contracts, exploring whether this is a strategic decision to support smaller businesses or a reflection of procurement policy regulations.

Refine Modeling Approaches: Advanced Statistical Models: Explore more complex modeling techniques, such as nonlinear regression, time series analysis, or machine learning models, to capture the intricate relationships between variables.

Incorporate Additional Variables: Include economic indicators, policy changes, and other relevant factors to improve the explanatory power of the models.

By undertaking these steps, future research can provide a more comprehensive and nuanced understanding of Canadian federal procurement spending. This further analysis could inform policy decisions, promote transparency, and ensure that government resources are allocated effectively and efficiently.

Appendix

A Web Scraping

These are step by step instructions on how to Web Scrape the IFJ procurement data page, because as of November 2024 there is no easy/automated way of downloading the 330 pages of data.

Note: It is important that docker, java, python and Firefox are installed on your machine as RSelenium requires these packages. To download Selenium and create a docker image, open the docker application and search for the Selenium package for the browser you wish to use, in this case Firefox.

1. Run a compatible Selenium docker image. You must make sure that the version of Selenium you are using is compatible with Firefox or whichever browser you choose. You must launch docker from your terminal whose directory matches that of the r file's path to 01-download_data.R . You can check if the docker image has successfully been launched using the docker application's nice GUI. It is also important in this step to define the path of the local directory that the data should be saved into, this is shown as USER_PATH in the script.

Terminal Command:

user:canadian_gov_procurement USER_PATH$ docker run -d -p 4444:4444 -p 5900:5900 -v /USER_PATH/canadian_gov_procurement/data/raw_data:/home/seluser/Downloads selenium/standalone-firefox:3.141.59

2. It is possible that the tags used to find the buttons on the web-site may change over time. So visit the IJF procurement data website and right click and inspect the download and next page buttons. This code uses the aria labels to locate these buttons, but there are many ways to identify them using html and css elements.

3. Now the 01-down-load-data.R file can be executed. Be sure to modify USER_PATH in both to download_dir and host_down_load_dir to match the path of you docker and personal computer path.

4. The for-loop simply runs "click download -> click next page" process 329 times, however IJF may add more pages so modify total_pages to scrape the number of pages IJF is offering. In the event of an error mid-way through scraping take note of which page failed to download (page that you are on is printed to the console) and simply set the start of the for-loop. After restarting the loop be sure to verify that the next downloaded page is the correct one, otherwise you will have to restart the process from step 1. While it may be tempting to reduce the Sys.sleep times to download faster (11 seconds delay per loop and 329 loops that's about 1 hour!), I found reducing the time cause errors as the pages were slow to load and thus ruining the process.

5. Done!

B Additional data details

TODO Removal of negative values, adding 1 day to contract lengths of 0, see cleaning script

C Proccessing Supplier

TODO processed_supplier Talk about what I did thourougly, see data_cleaning script

Talk about next steps. 11th top supplier, and how there are still tons of useless words in there. Even how to approach I used may never work. Because at some point we will ban too many word and some suppliers with super generic names will be eliminated entirely.

Propose alternate methods

D Model details

D.1 Diagnostics

TODO Add Diagnostics

E References

TODO Ref:

RSelenuim VGAM stringr dplyr tidyverse zoo kableExtra knitr scales lubridate ggplot2 rvest readr testthat