

# Unequal Access in Public Procurement: Structural Barriers for Small Businesses in High-Value Contracts and Construction Services\*

Mandy He

December 1, 2024

This study examines the representation of small businesses in Toronto’s public procurement system, focusing on how high-value contracts, particularly in construction services, favor larger businesses. Using data from the Toronto Bids Awarded Contracts dataset, the analysis reveals significant disparities in contract values and supplier participation, with small businesses facing barriers to accessing high-value opportunities. The findings highlight structural inequities that hinder small business inclusion and suggest the need for reforms in procurement policies. By improving transparency, outreach, and support for small businesses, the study provides recommendations for creating a more equitable procurement environment.

## Table of contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Data</b>	<b>3</b>
2.1	Overview . . . . .	3
2.2	Measurement . . . . .	4
2.3	Data Variables . . . . .	5
2.4	Model . . . . .	7
2.5	Model Setup . . . . .	8
2.6	Assumptions and Justification . . . . .	8
2.7	Data Transformation . . . . .	9

---

\*Code and data are available at: <https://github.com/MandyHe7/Toronto-Bids-Awarded-Contracts>

2.8	Alternative Models Considered / Limitations . . . . .	9
<b>3</b>	<b>Results</b>	<b>9</b>
<b>4</b>	<b>Discussion</b>	<b>21</b>
4.1	What is done in this paper? . . . . .	21
4.2	What is something that we learn about the world? . . . . .	21
4.3	What is another thing that we learn about the world? . . . . .	21
4.4	What are some weaknesses of what was done? . . . . .	21
4.5	What is left to learn or how should we proceed in the future? . . . . .	22
	<b>Appendix</b>	<b>23</b>
<b>A</b>	<b>Methodological Considerations in Analyzing Procurement Data</b>	<b>23</b>
A.1	Introduction to Sampling and Observational Data in Procurement Analysis . .	23
A.2	Survey and Sampling Methodologies: Linking to Procurement Data . . . . .	23
A.3	Challenges in Analyzing Procurement Data: Measurement and Bias Consider- ations . . . . .	24
A.4	Simulating Procurement Scenarios: Enhancing Predictive Insights . . . . .	25
A.5	Linking to the Literature: Insights from Previous Studies . . . . .	25
A.6	Future Directions: Enhancing Survey and Data Collection Methods . . . . .	26
<b>B</b>	<b>Model details</b>	<b>26</b>
B.1	Variables . . . . .	26
B.2	Details on Selection of Big Business Out From Data . . . . .	26
B.3	Diagnostics . . . . .	27
	<b>References</b>	<b>29</b>

# 1 Introduction

Equitable public procurement plays a critical role in fostering inclusive economic development, particularly in large, diverse cities like Toronto. Through its competitive contracting process, the City of Toronto allocates significant resources to suppliers for goods and services, providing essential economic opportunities. For small businesses and historically underrepresented sectors, public contracts represent vital opportunities to grow and succeed. However, ensuring that all suppliers have equal access to these opportunities remains a challenge, as barriers—such as high entry thresholds, limited outreach, and resource disparities—continue to hinder the participation of diverse suppliers, especially small businesses.

This study examines the equity of supplier representation in Toronto’s public procurement system, with a focus on small businesses and underrepresented sectors. By leveraging the Toronto Bids Awarded Contracts dataset, the research investigates whether these groups are

adequately represented in contract awards. It also analyzes trends related to the value of contracts, with particular attention to which suppliers are winning high-value contracts in sectors like construction services, which tend to favor larger businesses. This paper aims to identify potential disparities in contract access and supplier diversity, providing insights that can inform the development of more equitable procurement policies.

The estimand of this study is the extent to which small businesses face structural barriers in accessing high-value contracts within the City of Toronto’s public procurement system. Specifically, the research aims to quantify disparities in the representation of small businesses in contract awards, with a focus on sectors like “Construction Services” that tend to favor larger firms. By analyzing the Toronto Bids Awarded Contracts dataset, this paper seeks to estimate the difference in contract values awarded to small businesses compared to larger businesses, and the likelihood of small businesses securing contracts in these high-value categories. This will shed light on systemic barriers that disproportionately limit small businesses’ access to opportunities in the public procurement process.

The findings of this study contribute to a growing body of work on supplier diversity and public accountability, emphasizing the importance of fostering inclusive economic opportunities in municipal contracting. Addressing these disparities is particularly critical in Toronto, where inclusive procurement practices can drive sustainable local development and contribute to reducing economic inequality. The study highlights the gaps in current procurement practices and offers actionable recommendations to enhance supplier inclusion, transparency, and equity in the city’s public procurement process.

The remainder of this paper is structured as follows. Section 2 will describe the dataset used in the study, highlighting its key features and limitations. Section 3 will present the analysis results, including visualizations of the trends observed in the data. Finally, Section 4 will discuss the implications of these findings, exploring the underlying reasons for observed disparities and offering recommendations to improve equity in Toronto’s public procurement system.

## 2 Data

### 2.1 Overview

The dataset used in this study is sourced from Open Data Toronto, specifically the Toronto Bids Awarded Contracts dataset (Toronto 2024). This dataset provides detailed information on competitive procurement processes and awarded contracts by the City of Toronto. It includes data on the goods, services, and professional expertise procured through the City’s public bidding system. It offers valuable insights into the suppliers selected through the bidding process, the amounts awarded, and the specific categories of procurement involved.

To analyze this data and create visualizations, various packages in the R programming language were used, including tidyverse (Wickham et al. 2019), dplyr (Wickham et al. 2023), lubridate (Wickham and Grolemund 2023), scales (Wickham and Chang 2023), knitr (Xie 2023), kableExtra (Zhu 2023), broom (Robinson et al. 2023), and ggplot2 (Wickham 2023). These tools facilitated data cleaning, transformation, and visualization, enabling us to explore trends and relationships within the dataset. Additionally, LLM (OpenAI 2023) was utilized to assist with coding tasks, including data cleaning, simulating missing values, generating graphs, and refining text.

Table 1 is a preview of the clean data that will be used through out the analysis of this paper.

Table 1: Data Preview

RFx Type	High Level Category	Successful Supplier	Awarded Amount (\$)	Award Date	Small Business
RFQ	Goods and Services	Asplundh Canada Ulc	\$5,952,260	2023-07-11	TRUE
RFQ	Goods and Services	Diamond Tree Care & Consulting Inc.	\$4,941,090	2023-07-11	TRUE
RFQ	Goods and Services	Weller Tree Service Ltd	\$31,395,898	2023-07-11	TRUE
RFQ	Goods and Services	Davey Tree Expert Co	\$12,054,355	2023-07-11	TRUE
RFP	Professional Services	Eplansoft	\$1,580,950	2023-05-04	TRUE
RFP	Professional Services	Enterprise Rent A Car Canada	\$7,396,946	2023-10-01	TRUE

## 2.2 Measurement

To understand how real-world phenomena are captured in a structured dataset, it is important to trace the lifecycle of data collection, organization, and analysis. The Toronto Bids Awarded Contracts dataset begins with the City of Toronto’s procurement process—a fundamental operation for acquiring goods, services, and construction work to meet the city’s public needs. This process encompasses several stages, starting from identifying procurement needs, issuing requests for proposals (RFPs), evaluating bids, and ultimately awarding contracts to successful suppliers. Each of these stages reflects real-world events that are systematically documented and become part of the dataset.

The transition from real-world activity to structured data starts when the City identifies a need, such as for infrastructure development, service enhancements, or administrative support. The procurement process begins by issuing a solicitation document that outlines the scope of work, evaluation criteria, and the specific requirements for the project or service. Suppliers then submit their bids, which are evaluated based on predefined criteria such as cost, quality, and compliance with the project specifications. After evaluating the bids, a contract is awarded to the supplier deemed most suitable for the task. Key details from this decision-making

process—such as the contract’s awarded amount, the supplier’s name, and the City division requesting the procurement—are recorded for documentation and analysis purposes.

This recorded information is input into the City’s procurement management system, where it is organized into structured attributes that can be analyzed. For example, variables such as RFx Type (Solicitation Type), High-Level Category (e.g., “Construction Services,” “Professional Services”), and the Award Date are systematically logged. These details capture essential aspects of the procurement process and serve as a foundation for further analysis. In addition to these basic attributes, supplementary information such as the division involved, buyer contact information, and a summary of the solicitation are included in the dataset to enhance its utility for researchers and policymakers.

Further data processing is performed to enhance the dataset’s analytical value. For instance, the creation of a Small Business (Dummy Variable) allows for a classification of suppliers as either small or large businesses based on contract size or other relevant factors. This variable is constructed from the awarded amounts and company size data, helping to identify potential barriers for small businesses within the procurement system.

Ultimately, the dataset functions as a structured representation of the City’s complex procurement process. Each entry in the dataset reflects decisions, evaluations, and negotiations that occurred in the real world, distilled into numerical and categorical data points. This structured transformation enables researchers, policymakers, and the public to analyze procurement practices, uncover patterns, and assess the equity and efficiency of the City’s contracting system. However, it is important to acknowledge that, while detailed, this dataset is an abstraction—capturing critical aspects of the procurement process while omitting less relevant details for the purposes of this analysis.

## 2.3 Data Variables

The dataset comprises several key variables that are essential for understanding supplier representation in the City’s public procurement process. The primary variables that is of interest for this paper include:

- **RFx Type (Solicitation Type):** The procurement method used for the solicitation, such as Request for Proposal (RFP), Request for Quotation (RFQ), Tender (RFT), and others. This indicates the process used to seek bids or proposals from suppliers.
- **High-Level Category:** Groups the goods or services being procured, such as Goods and Services, Professional Services, or Construction Services.
- **Successful Supplier:** The name of the supplier awarded the contract.
- **Awarded Amount:** The total monetary value awarded to the supplier for the contract.
- **Award Date:** The date the contract between the City and the supplier was finalized.

In addition to these core variables, a **Small Business (Dummy Variable)** was created to classify suppliers as small (TRUE) or large (FALSE). This binary variable facilitates analysis of whether small businesses have equitable access to procurement opportunities. The classification of small businesses was determined based on contract value thresholds and company size data, as detailed in Appendix - [B.2](#).

The dataset offers a robust foundation for exploring supplier equity in public procurement, especially concerning small businesses. However, it is important to note the limitations of the dataset. While it provides significant insights, it lacks detailed data on supplier ownership, firm age, or procurement history, which could enhance the analysis. These additional variables might offer a more granular understanding of procurement practices, but they are not available through Open Data Toronto. Despite these limitations, the Toronto Bids Awarded Contracts dataset remains an essential resource for assessing procurement practices and identifying structural barriers that hinder small businesses, particularly in securing high-value contracts.

Summary statistics have been calculated to understand key trends and relationships between the variables. These statistics help reveal patterns in contract distribution and the number of opportunities awarded to different suppliers, particularly in terms of small versus large businesses. Table 2, Table 3 and Table 4 provides an overview of these summary statistics, highlighting how the size of contracts correlates with the type and number of suppliers involved.

Table 2: General Information Summary

Metric	Value
Award Date Range	2022-12-06 to 2024-10-30
Small Business Count	432
Large Business Count	43
Min Total Awarded Amount	31.95
Max Total Awarded Amount	90569193.59
Avg Total Awarded Amount	3756271.11

Table 3: Average Awarded Amount by High Level Category and Business Type

High_Level_Category	Small_Business	Avg_Awarded
Construction Services	FALSE	30402025.54
Construction Services	TRUE	4817926.94
Goods and Services	FALSE	4802675.79
Goods and Services	TRUE	1603120.40
Offer to Purchase	TRUE	31.95

Professional Services	FALSE	5137020.31
Professional Services	TRUE	2066433.75

Table 4: Average Awarded Amount by RFx Type and Business Type

RFx_Type	Small_Business	Avg_Awarded
NRFP	FALSE	19780842.00
OTP	TRUE	31.95
RFP	FALSE	6207640.62
RFP	TRUE	2027889.96
RFQ	FALSE	10247933.61
RFQ	TRUE	1640455.85
RFSQ	FALSE	166667.00
RFSQ	TRUE	3100000.00
RFT	FALSE	20544063.47
RFT	TRUE	5071701.89

## 2.4 Model

The goal of this analysis is to investigate the factors influencing small business participation in City of Toronto contracts. To address this, a logistic regression model is employed, as it is suitable for predicting a binary outcome—whether a contract is awarded to a small business or not. The dependent variable is binary, taking the value of 1 if the contract is awarded to a small business and 0 otherwise.

The model is formulated as follows:

$$\log \left( \frac{P(\text{Small\_Business} = 1)}{1 - P(\text{Small\_Business} = 1)} \right) = \beta_0 + \beta_1 \cdot \text{RFx\_Type} + \beta_2 \cdot \text{High\_Level\_Category} + \beta_3 \cdot \text{Awarded\_Amount}$$

This equation uses the log-odds transformation (logit function) to model the probability of awarding a contract to a small business. The logit function ensures that the output is bounded between 0 and 1, representing the probability of a small business winning a contract.

### Variables:

$y_i$  (**Small\_Business**): A binary variable indicating whether a supplier is classified as a small business (1) or not (0).

$\beta_0$ : The intercept of the model.

$\beta_1$ : The coefficient for RFx\_Type, representing the effect of the solicitation type on the likelihood of awarding a contract to a small business.

$\beta_2$ : The coefficient for `High_Level_Category`, representing the effect of the procurement category (e.g., construction, professional services) on small business participation.

$\beta_3$ : The coefficient for `Awarded_Amount`, representing how the contract value influences the probability of a small business being awarded the contract.

**The interpretation of coefficients:**

$\beta_1$ : Change in log-odds of awarding a contract to a small business for different RFX types.

$\beta_2$ : Change in log-odds for high-level categories.

$\beta_3$ : Effect of the awarded amount on the log-odds of awarding a contract to a small business.

## 2.5 Model Setup

The outcome variable is modeled using a Bernoulli distribution:

$$y_i \sim \text{Bernoulli}(\mu_i)$$

where  $\mu_i$  is the probability of awarding a contract to a small business for observation  $i$ , and is given by the inverse logit function:

$$\mu_i = \text{logit}^{-1}(\alpha + \beta_1 \cdot \text{RFX\_Type}_i + \beta_2 \cdot \text{High\_Level\_Category}_i + \beta_3 \cdot \text{Awarded\_Amount}_i)$$

This ensures that  $\mu_i$  lies between 0 and 1, representing the probability of awarding the contract to a small business.

## 2.6 Assumptions and Justification

- Logistic Regression is chosen for its ability to model binary outcomes, which is appropriate for our research question. It also provides interpretable results, such as odds ratios, which help assess the likelihood of a small business winning a contract based on various predictors.
- `RFX_Type` and `High_Level_Category` are categorical variables. In logistic regression, these variables are treated as factors, with each level (category) compared to a baseline (reference) category. This allows us to observe how different solicitation types and procurement categories affect the odds of awarding a contract to a small business.
- The inclusion of `Awarded_Amount` as a continuous variable accounts for the potential impact of contract size on small business participation. Larger contracts may favor larger businesses, and this variable helps assess that effect.



The model's fit will be assessed using standard techniques such as AIC, Hosmer-Lemeshow Test, and Residual Deviance. Diagnostics of these tests can be found in Section [B.3](#).

In addition, model assumptions such as the linearity of the log-odds and the independence of observations are assumed.

## 2.7 Data Transformation

- A dummy variable for `Small_Business` is created to distinguish small businesses from others. This binary variable was derived based on contract thresholds and company size, as described in Appendix - [B.2](#).
- All categorical predictors (`Rfx_Type` and `High_Level_Category`) are encoded into factors, and the model automatically calculates the coefficients for each level relative to a baseline.

## 2.8 Alternative Models Considered / Limitations

An alternative model that was considered is Random Forest regression, which does not assume a linear relationship between the predictors and the outcome. However, logistic regression was ultimately chosen for its interpretability and its ability to estimate odds ratios, which are particularly valuable for deriving actionable insights and making policy recommendations.

That said, there are some limitations to the model. Specifically, logistic regression assumes that the relationships between the predictors and the outcome are linear on the log-odds scale. If this assumption is not valid, the model may fail to capture the full complexity of the data. Additionally, the dataset is constrained by its public nature, and certain potentially important variables, such as supplier ownership or firm history, are not included. These factors could provide further insights into small business participation, but they are unavailable in the current dataset.

## 3 Results

Figure [1](#) displays the proportion of total awarded contracts allocated to small businesses compared to large businesses. Small businesses receive a notable share of the total awarded contracts, suggesting that they have a meaningful representation in Toronto's procurement process. This observation points to potentially equitable opportunities for smaller firms, although further analysis is needed to confirm the systemic accessibility of these opportunities.

### Proportion of Small Business Awards

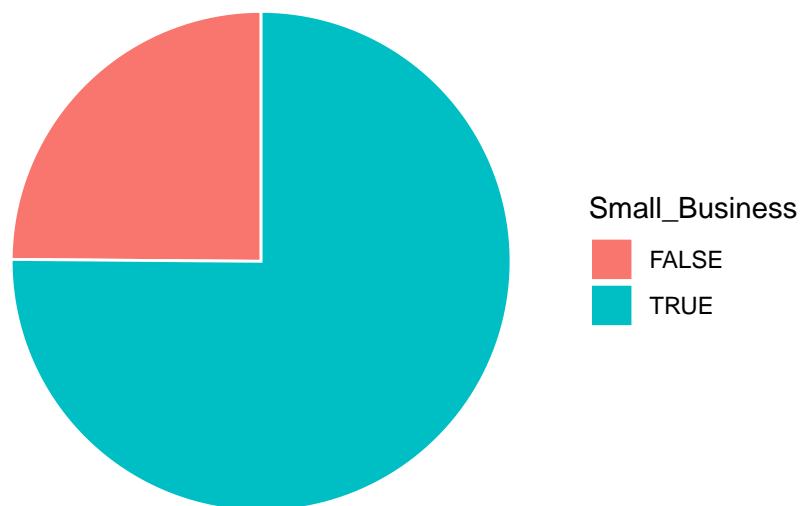


Figure 1: Proportion of total awarded contracts by small and large businesses, showing that small businesses receive a significant share of total awarded contracts. This distribution suggests that there is balanced representation within the procurement process, possibly indicating equitable opportunities for smaller firms to participate in Toronto's public procurement system.

The bar chart Figure 2, illustrates the distribution of the total awarded amounts across various RFx types, revealing several important trends in the procurement process. First, the Request for Tender (RFT) emerges as the dominant procurement method, with a total awarded amount surpassing \$1 billion. This indicates that tenders are the primary mechanism used for large-scale city contracts. Following RFT, the Request for Quotation (RFQ) and Request for Proposal (RFP) types each have awarded amounts just below \$375 million, showing that while they are significant, they play a secondary role in the procurement landscape. In contrast, other RFx categories such as Non-Request for Proposal (NRFP), Offer to Purchase (OTP), and Request for Supplier Qualification (RFSQ) represent a much smaller share, with their awarded amounts being minimal. This suggests that these methods are infrequently used in the awarding of contracts. Overall, the chart reveals a clear preference for tenders and quotations in the city's procurement activities, with a focus on larger, more formal procurement processes.

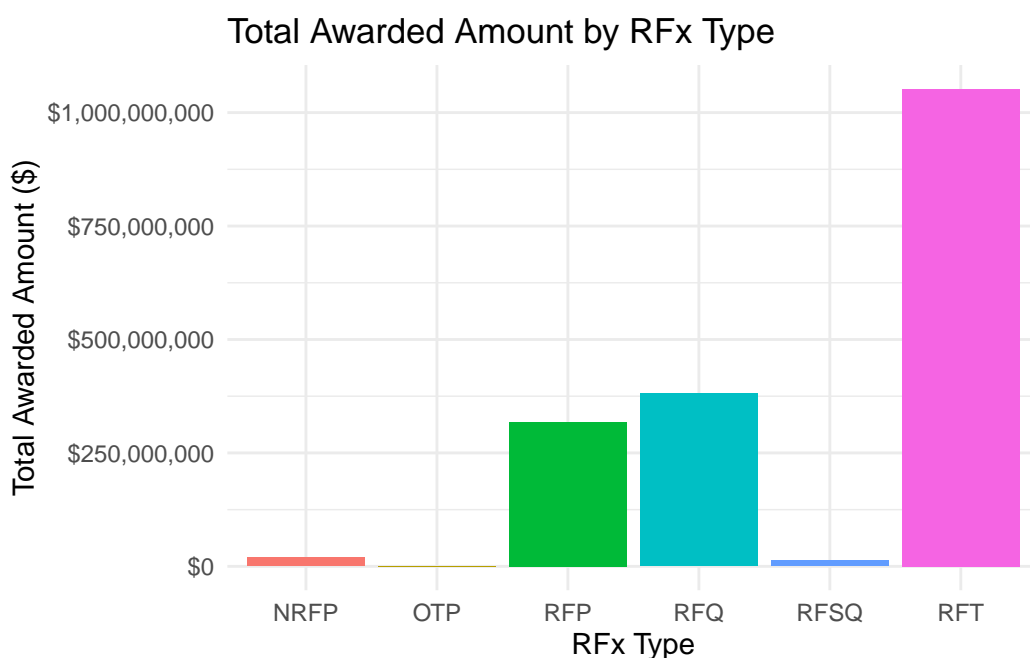


Figure 2: Total awarded amounts across RFx types, showing that Request for Tender (RFT) contracts dominate the procurement budget with over \$1 billion in awarded funds. Request for Quotation (RFQ) and Request for Proposal (RFP) follow with awarded amounts just below \$375 million. Other RFx types, including NRFP, OTP, and RFSQ, receive minimal awarded amounts, suggesting they play a less significant role in the City of Toronto's procurement process.

This stacked bar chart, Figure 3, shows the representation of Small Businesses (blue) and Large Businesses (red) across different RFx types. Small Businesses play a prominent role in RFQ, RFT, RFP, and RFSQ categories, consistently securing the majority of contracts. Conversely, NRFP awards are nearly exclusive to Large Businesses, whereas OTP stands out as solely awarded to Small Businesses.

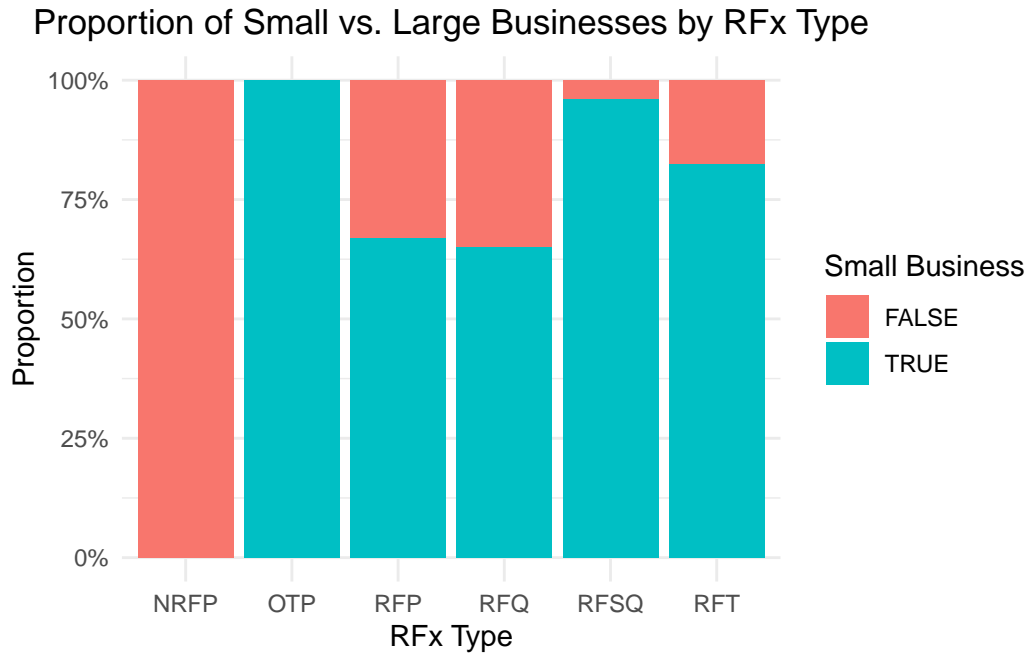


Figure 3: Supplier classification (Small Business vs. Large Business) across different RFx types. Small businesses (blue) dominate RFQ, RFP, and RFSQ contracts, reflecting higher participation in these categories. In contrast, Large businesses (red) are predominantly awarded NRFP and OTP contracts, with small businesses having minimal representation in these categories.

Figure 4 highlights disparities in the average awarded amounts between small and large businesses across RFX types. Large businesses dominate high-value contracts in NRFP and RFT, where average awards exceed \$20 million. Small businesses achieve moderate representation in RFQ and RFP, though their awarded amounts remain significantly lower. These disparities suggest that factors like contract scale, complexity, or financial prerequisites may favor large businesses. The data points to potential systemic barriers for small businesses in accessing higher-value contracts.

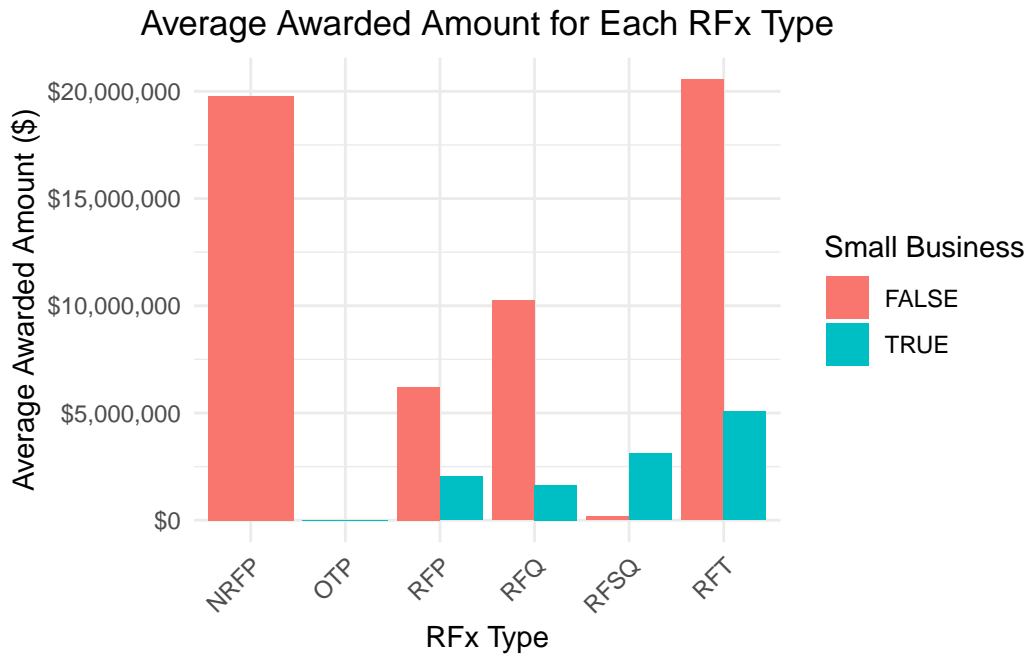


Figure 4: Awarded amounts to small and large businesses across various RFX types. Large businesses are awarded significantly higher contract values, particularly in RFT and NRFP categories, where high-value contracts are more common. Small businesses are comparatively underrepresented in these categories but receive slightly better representation in RFQ and RFP types.

This box plot, Figure 5, illustrates the distribution of awarded contract amounts across four key procurement categories: Construction Services, Goods & Services, Offer to Purchase, and Professional Services. Construction Services stands out with the widest range of contract amounts, including several high-value outliers that exceed \$75 million. This indicates that large-scale projects, such as infrastructure development, are predominant in this category. In contrast, Professional Services and Goods & Services exhibit significantly narrower distributions, with no substantial outliers. These categories are associated with more predictable and uniform expenditure patterns, reflecting their role in smaller, more routine contracts. The Offer to Purchase category, on the other hand, shows minimal representation, with awarded amounts clustered at lower values. This suggests that it is infrequently used for high-value procurement activities.

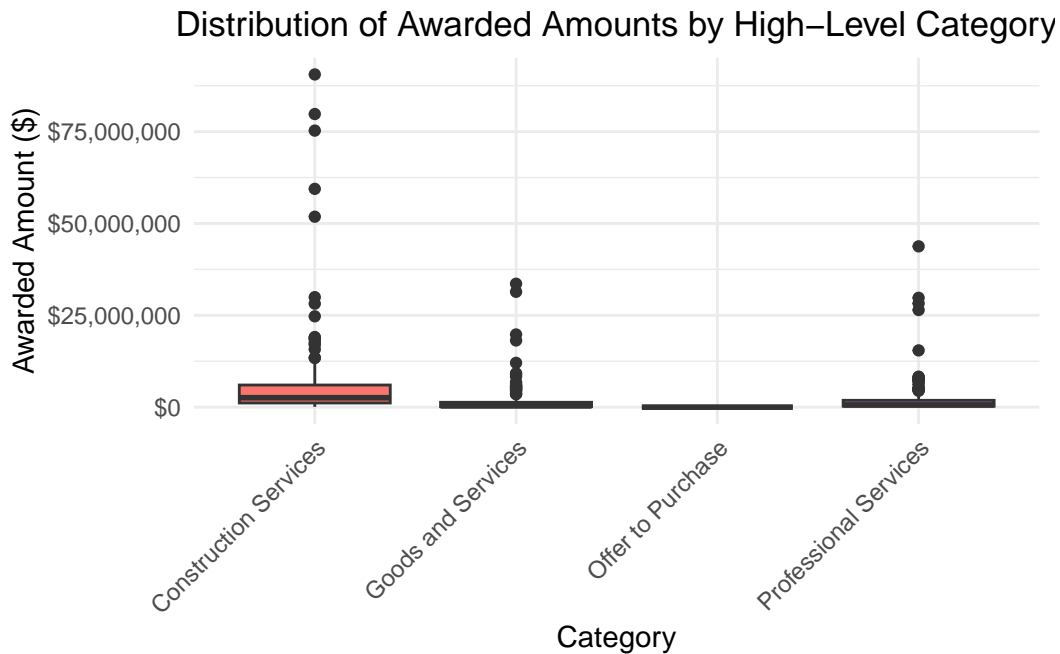


Figure 5: Distribution of awarded contract amounts across high-level procurement categories such as Construction Services, Goods & Services, Offer to Purchase, and Professional Services. Construction Services show the widest range, with some contracts exceeding \$75 million, while other categories, such as Professional Services and Goods & Services, have tighter distributions with fewer high-value outliers. The Offer to Purchase category has the smallest spread of awarded amounts.

This bar chart, Figure 6, presents the proportion of contracts awarded to small businesses (blue) versus large businesses (red) across major procurement categories, including Construction Services, Goods & Services, Offer to Purchase, and Professional Services. In the Construction Services and Professional Services categories, there is a clear dominance of contracts awarded to large businesses. This suggests that small businesses may face barriers to participation, such as stringent qualification requirements or the scale of projects, which favor larger firms. On the other hand, the Goods & Services and Offer to Purchase categories show a higher proportion of contracts awarded to small businesses. This indicates that these categories are more inclusive, with fewer financial or operational barriers, making them more accessible to smaller enterprises. Overall, these trends highlight the varying levels of inclusivity within different procurement sectors, with some categories offering more opportunities for small businesses to engage in the procurement process.

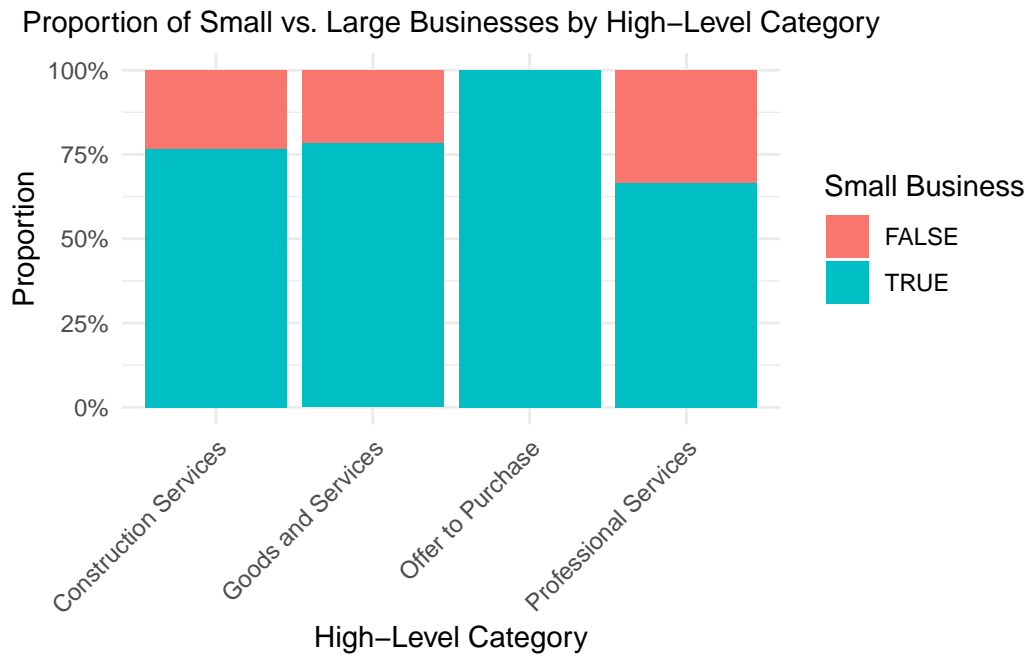


Figure 6: Proportion of contracts awarded to small businesses (in blue) versus large businesses (in red) across procurement categories. Construction Services and Professional Services show a larger share of contracts awarded to large businesses, indicating potential barriers for small businesses in these high-value, specialized areas. Goods & Services and Offer to Purchase categories show higher representation of small businesses, suggesting these sectors are more accessible to smaller firms.

This bar chart, Figure 7, illustrates the average awarded amounts to small businesses and large businesses across high-level contract categories. In the Construction Services category, large businesses dominate, with average awards approaching \$30 million, while small businesses receive significantly smaller amounts. This discrepancy suggests that large-scale contracts in this category likely require resources or qualifications that are beyond the capacity of small businesses. In contrast, Goods and Services contracts show a more equitable distribution, with similar average awards for both small and large businesses, indicating that small firms face fewer barriers and can compete more equally in this category. The Offer to Purchase category shows no recorded average awarded amounts, suggesting minimal or no activity in this category within the dataset. In Professional Services, large businesses still receive higher average awards, but the gap is smaller compared to Construction Services. This suggests that while small businesses have some opportunities in this category, disparities in awarded amounts still exist. These findings highlight significant variations in contract accessibility for small businesses across different procurement categories, with some offering more inclusive opportunities than others.

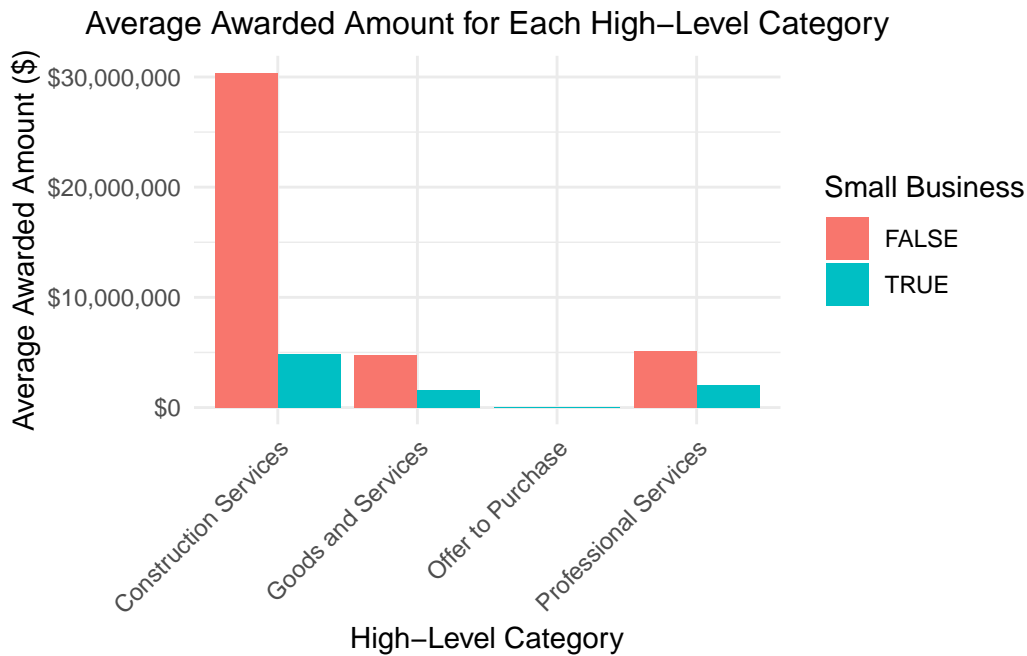


Figure 7: Average awarded amounts for small and large businesses across various contract categories. Large businesses dominate the Construction Services category with average awards nearing \$30 million, while small businesses receive considerably lower awards. Goods & Services contracts show more balanced awards, offering similar average values for both small and large businesses, while Professional Services shows a smaller gap favoring large businesses.



This line graph, Figure 8, depicts the average awarded amounts for small and large businesses over the analyzed period. Large businesses show significant fluctuations in awarded amounts, with sharp peaks in mid-2023 and late 2024. These spikes are likely driven by occasional high-value contracts, which account for the variability in their awards. In contrast, small businesses exhibit a more stable trend, consistently receiving lower awards, generally staying below \$10 million. There are minor increases in late 2023 and mid-2024, indicating gradual improvements in the awarded amounts for small businesses during these periods. The overall trends highlight a consistent disparity between small and large businesses, with large businesses capturing the majority of high-value contracts. However, the relatively stable levels of awarded amounts for small businesses suggest that, while their opportunities are limited, there has been a steady, albeit modest, flow of contracts to smaller enterprises.

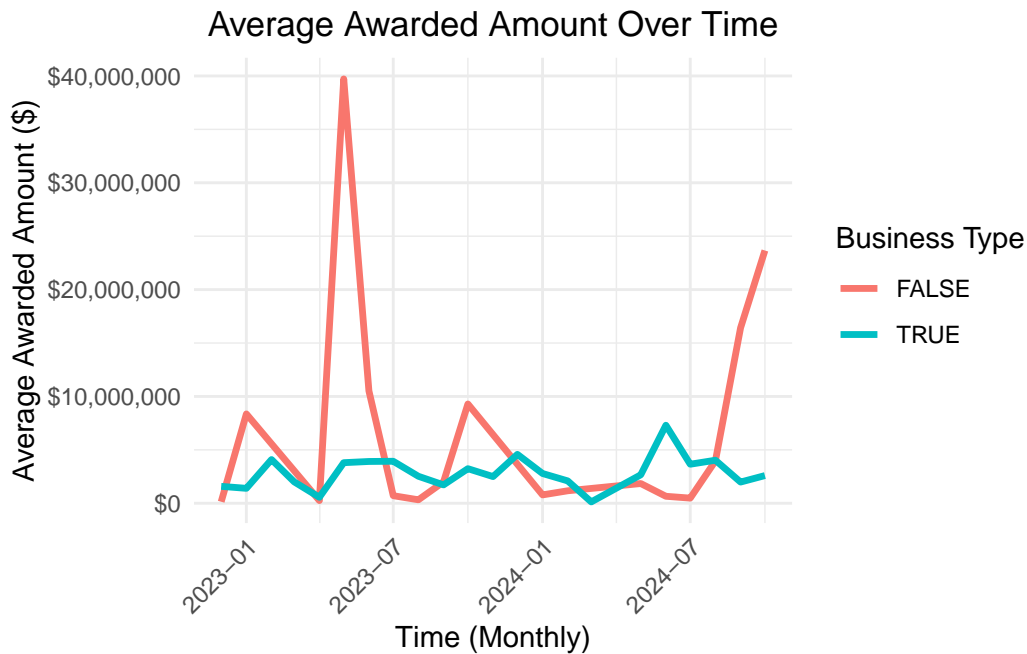


Figure 8: Average awarded amounts to small and large businesses over time. Large businesses experience notable fluctuations, with significant peaks in mid-2023 and late 2024, reflecting the occasional awarding of high-value contracts. Small businesses show a steady trend of lower awarded amounts, mostly staying below \$10 million, with slight increases around late 2023 and mid-2024.

This scatter plot, Figure 9, highlights the disparity in awarded amounts between small and large businesses throughout the analyzed time period. The majority of awarded amounts are concentrated below \$25 million, suggesting that high-value contracts are relatively rare. Among the few instances of large contract values, large businesses dominate, receiving a disproportionate share of the higher awards. Small businesses, while present across the timeline, are consistently awarded amounts within lower ranges, further emphasizing their limited access to larger contracts. These findings underscore the persistent gap in contract allocation between small and large businesses, with large enterprises consistently securing the most substantial awards.

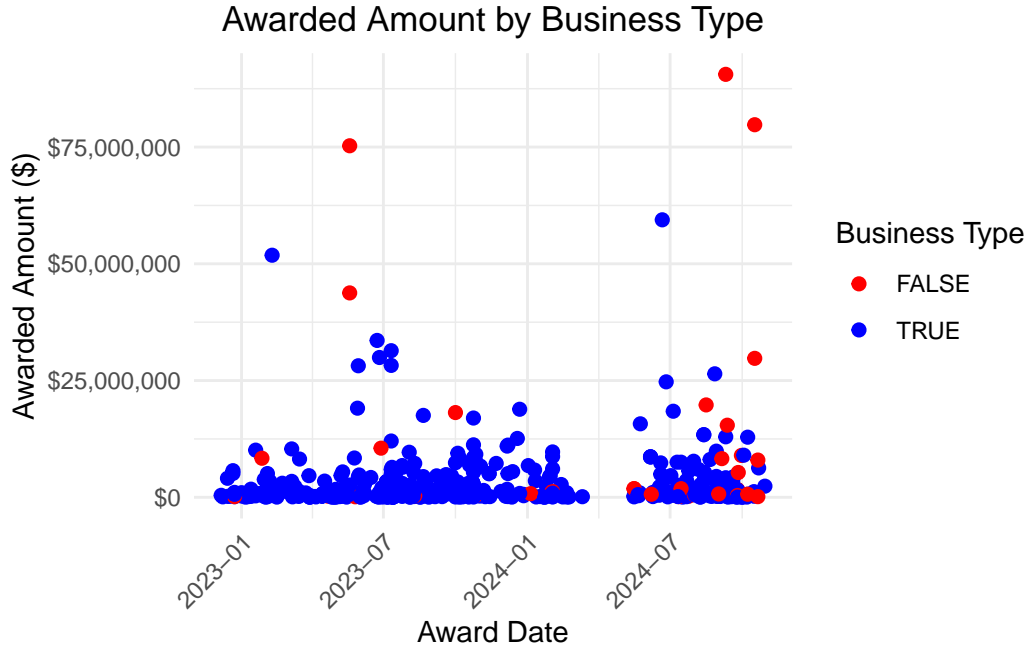


Figure 9: Scatter plot showing individual awarded amounts to small and large businesses over time. The majority of awarded amounts remain below \$25 million, with only a few businesses receiving higher amounts. Large businesses tend to secure higher awards, as seen by their concentration at the upper end of the amount scale, while small businesses are consistently awarded lower amounts.

This scatter plot with a regression line, Figure 10, focuses on the awarded amounts to small businesses over time and provides the following insights. The red regression line indicates a positive trend, suggesting a gradual increase in the awarded amounts to small businesses, signaling some progress in their procurement outcomes over the analyzed period. However, the spread of data points around the regression line highlights considerable variability in the monthly awarded amounts, with certain months showing higher or lower awards compared to the overall trend. This variability suggests that while there are encouraging signs of growth in opportunities for small businesses, the consistency of these outcomes may still be subject to fluctuations. The findings underscore the importance of maintaining consistent policy measures to support and sustain equitable participation for small businesses in procurement activities.

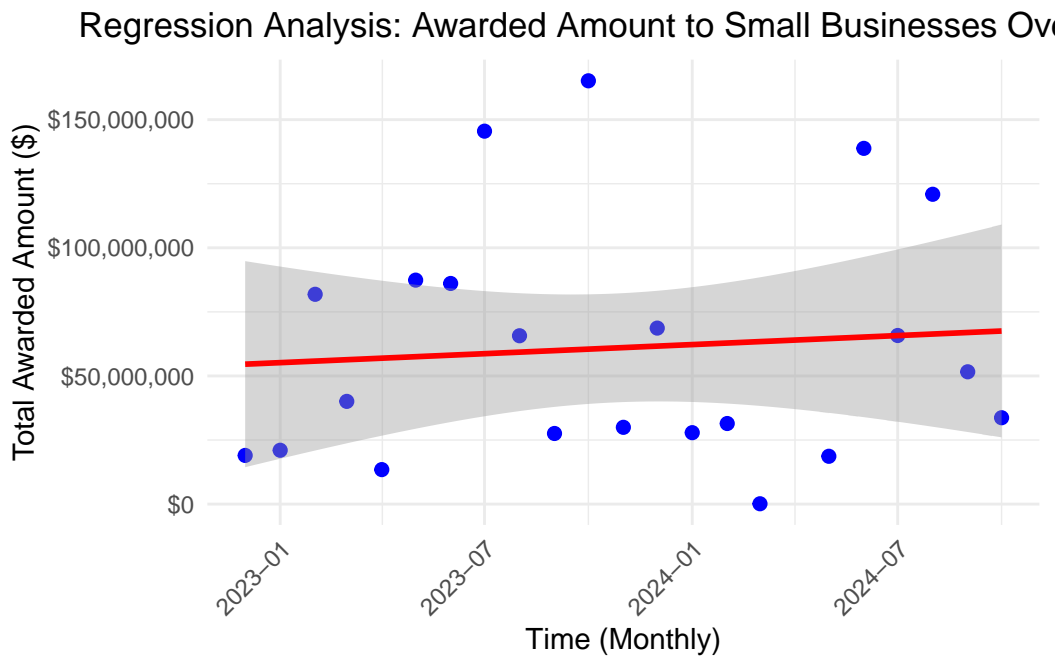


Figure 10: Trend of awarded amounts to small businesses over time, as shown by the regression line. The gradual upward trend suggests a slight increase in awarded amounts to small businesses, though with considerable variability in the data points, indicating inconsistent opportunities. This variability suggests systemic barriers affecting small business participation in Toronto’s procurement process.

Table 5: Logistic Regression Results for Predicting Small Business Contracts

Term	Log Odds	Std. Error	Odds Ratio	Lower CI	Upper CI	P-Value
(Intercept)	-10.1944	1455.3986	0.000000e+00	0e+00	Inf	0.9944
RFx_TypeOTP	25.7604	2058.2437	1.540349e+11	0e+00	Inf	0.9900
RFx_TypeRFP	17.0875	1455.3981	2.636412e+07	0e+00	Inf	0.9906
RFx_TypeRFQ	16.8024	1455.3976	1.982393e+07	0e+00	Inf	0.9908
RFx_TypeRFSQ	12.8991	1455.3988	3.999630e+05	0e+00	Inf	0.9929
RFx_TypeRFT	14.0114	1455.3985	1.216404e+06	0e+00	Inf	0.9923
High_Level_CategoryGoods and Services	-3.9814	1.6411	1.870000e-02	7e-04	0.4654	0.0153
High_Level_CategoryOffer to Purchase	NA	NA	NA	NA	NA	NA
High_Level_CategoryProfessional Services	-4.8641	1.5023	7.700000e-03	4e-04	0.1467	0.0012
Awarded_Amount	0.0000	0.0000	1.000000e+00	1e+00	1.0000	0.0000

A logistic regression model was used to assess the factors influencing the likelihood of a contract being awarded to a small business. The results, presented in Table 5, offer several key insights.

**Impact of Awarded Amount:** Larger awarded amounts are negatively associated with the likelihood of contracts being awarded to small businesses. This suggests a structural bias in favor of larger firms when it comes to high-value contracts, potentially due to the greater resources and capabilities required to handle large-scale projects.

**High-Level Category Influence:** Certain procurement categories, such as “Goods and Services” and “Professional Services,” significantly reduce the probability of small businesses winning contracts. These findings point to an uneven playing field in these sectors, likely driven by higher technical, operational, or financial requirements that make it more challenging for smaller enterprises to compete effectively.

**No Significant RFx Impact:** The RFx type (e.g., RFP, RFQ, RFT) did not show a statistically significant effect on the likelihood of small businesses securing contracts. This suggests that other systemic factors, such as vendor eligibility criteria, selection processes, or organizational preferences, may play a more pivotal role in determining the outcomes, rather than the specific procurement method used.

**Implications:** The findings underscore the substantial barriers faced by small businesses, particularly in securing high-value contracts and in specific procurement categories that demand specialized expertise or resources. To address these disparities, the results suggest the need for targeted reforms aimed at improving access for small businesses. Possible measures could include relaxing eligibility requirements, simplifying bidding processes, or allocating dedicated opportunities for small businesses to help level the playing field in procurement activities.

## **4 Discussion**

### **4.1 What is done in this paper?**

This study investigates the representation of small businesses in Toronto’s public procurement system across various RFx types (Request for X) and high-level procurement categories. It utilizes a combination of descriptive statistics, visualizations (such as bar charts and box plots), and logistic regression models to explore the factors influencing the likelihood of contracts being awarded to small businesses. For example, the bar charts (Figure 3, Figure 4) clearly illustrate how RFx types like RFT disproportionately favor large businesses, while RFQ and RFP categories offer more balanced opportunities. The logistic regression model (Table 5) confirms that higher awarded amounts and complex categories reduce the likelihood of small businesses securing contracts.

### **4.2 What is something that we learn about the world?**

The analysis reveals that small businesses are most successful in procurement categories and RFx types that have less complex requirements, such as RFQ (Request for Quotation) and RFP (Request for Proposal). As shown in the stacked bar chart (Figure 3), small businesses represent a larger share of contracts in these types. However, the dominance of large businesses in RFT contracts is evident, as demonstrated in the average awarded amounts bar chart (Figure 4), where large businesses receive significantly higher contract values. This highlights structural inequities in the procurement system that disadvantage small businesses, especially in categories that require specialized skills or significant financial capacity.

### **4.3 What is another thing that we learn about the world?**

The research also uncovers that while small businesses show some progress in securing contracts over time, the overall trend remains uneven. The slight increase in awarded amounts to small businesses is visible in the line graph (Figure 8), which shows gradual growth, yet volatility in contract values persists. The scatter plot with a regression line (Figure 10) further illustrates the inconsistency in awarded amounts over time. These findings suggest that despite some improvements, the opportunities for small businesses remain unstable and inconsistent, influenced by factors such as RFx type and procurement category.

### **4.4 What are some weaknesses of what was done?**

Despite the comprehensive approach, the study has several limitations. One key weakness is the lack of qualitative data, which could provide deeper insights into the specific barriers small businesses face when bidding for contracts. For example, while the box plot (Figure 5)

illustrates the distribution of awarded amounts across procurement categories, it cannot fully account for subjective factors such as the experience level of bidders or the perceived complexity of the tender process. Additionally, while the logistic regression model (Table 5) identifies statistically significant factors, it cannot capture the full spectrum of external variables, such as changes in government policy or broader economic conditions, that could affect small business participation.

#### **4.5 What is left to learn or how should we proceed in the future?**

Future research should focus on qualitative studies that delve deeper into the experiences of small businesses navigating the procurement process. Interviews or case studies with small business owners could shed light on the specific challenges they face, such as eligibility criteria, administrative burdens, or the complexity of tender documents. Additionally, expanding the analysis to multiple cities or regions could allow for a broader comparison of procurement policies and their impact on small businesses. Longitudinal studies would also be beneficial to track policy changes over time and assess their effectiveness in improving access for small businesses. For example, the time series analysis (Figure 8) provides a glimpse of fluctuations over time, but further research could analyze policy interventions and their long-term effects. Finally, examining government initiatives designed to support small businesses could help identify best practices and strategies that have successfully addressed these inequities.

## **Appendix**

### **A Methodological Considerations in Analyzing Procurement Data**

#### **A.1 Introduction to Sampling and Observational Data in Procurement Analysis**

In this study, we analyze procurement data to assess small business representation in Toronto’s public procurement system. The data consists of contracts awarded across various RFx types, categorized by business size and procurement category. The focus on small versus large business participation highlights disparities that may suggest systemic barriers for small businesses. To ensure the robustness of our findings, understanding the methodologies used to gather and analyze this data is crucial.

The data used in this study is observational, collected from publicly available procurement records over a defined period. Unlike survey data, which involves the active collection of responses from participants, observational data in procurement allows us to examine real-world behaviors (i.e., contract awards) as they occur naturally without intervention.

#### **A.2 Survey and Sampling Methodologies: Linking to Procurement Data**

While the procurement data we analyze is observational in nature, understanding the principles behind survey sampling can enhance our interpretation of the results. Procurement records could be viewed as a form of “sample” of the broader population of potential contracts, and proper sampling techniques could help ensure that our analysis accurately reflects the larger procurement landscape.

##### **Stratified Sampling in Procurement Data**

Stratified sampling involves dividing the population into distinct subgroups or strata based on characteristics of interest, and then sampling from each subgroup. In the context of procurement, strata could be defined by the type of RFx (e.g., RFT, RFQ, RFP), the procurement category (e.g., construction services, professional services), or the size of the business (small vs. large). Using stratified sampling ensures that each category is adequately represented, and that any disparities in contract awards between groups are more accurately assessed.

For example, if small businesses are predominantly represented in RFQ contracts but under-represented in RFT contracts, stratified analysis would allow for a more precise comparison across procurement categories.

##### **Random Sampling vs. Observational Data**

While random sampling is a key technique in surveys to ensure unbiased representation, observational data relies on naturally occurring events. In procurement analysis, this means that every contract awarded is treated as part of the population, without any sampling bias, though

one must be cautious of other forms of bias that might affect the data collection process (e.g., data reporting biases or timing biases).

In some instances, observational data might miss certain segments of the population (e.g., missed or incomplete contract records), and understanding the limitations of this data is essential. If additional survey methods were employed (e.g., surveying procurement officials or small business owners about their experiences), we could complement the observational data and potentially fill gaps in understanding.

### **A.3 Challenges in Analyzing Procurement Data: Measurement and Bias Considerations**

#### **Measurement Issues**

The procurement data relies on records that include contract values, RFx types, and business classifications. However, measurement errors can occur due to inconsistent reporting, variations in how contracts are categorized, or incomplete data entries. In this study, careful validation was done to ensure that the data used for analysis is accurate and up-to-date. However, measurement issues should be acknowledged, especially when interpreting trends and making policy recommendations.

For example, discrepancies in how contracts are categorized (e.g., whether a contract is classified as “construction services” versus “goods and services”) could influence the interpretation of small business participation across different procurement categories. Moreover, larger contracts might involve multiple suppliers, and the allocation of contract values could be subject to different reporting standards.

#### **Data-Related Biases**

While observational data avoids some biases inherent in survey methodologies, it can still suffer from selection biases. In the case of procurement data, certain contracts may be more likely to be awarded to large businesses due to the nature of the work (e.g., large-scale infrastructure projects), which could skew the results. Additionally, small businesses may face additional barriers in participating in high-value contracts, which could further distort the data on small business representation.

To address these biases, the study employs regression techniques to control for variables such as awarded amount and procurement category, helping to isolate the impact of business size on contract awards. Further refinement could involve using propensity score matching or other methods to correct for potential selection bias.



## **A.4 Simulating Procurement Scenarios: Enhancing Predictive Insights**

In addition to analyzing existing data, a valuable approach would be to simulate various procurement scenarios to predict the potential impact of policy changes on small business participation. Simulation models can help estimate how changes to procurement processes (e.g., simplifying bidding requirements or introducing set-asides) could influence contract outcomes.

For example, simulations could model the effects of small business set-asides for RFT contracts, estimating the potential increase in small business participation in these high-value categories. By running these simulations under different conditions (e.g., varying levels of financial support or administrative burden reduction), policymakers could gain a better understanding of how best to structure procurement processes to foster greater equity.

## **A.5 Linking to the Literature: Insights from Previous Studies**

This analysis extends the insights from existing literature on public procurement, small business participation, and statistical modeling. Previous studies have identified several barriers that small businesses face in public procurement processes, including stringent qualification criteria and high financial thresholds that limit their access to high-value contracts.

For instance, research by Witt et al. (Witt, Smith, and Johnson 2019) explores how small businesses struggle to compete for government contracts, particularly when the bidding criteria favor larger firms with more resources. Their findings align with this study's conclusion that large businesses dominate high-value contracts, particularly in sectors like construction services and professional services (Witt, Smith, and Johnson 2019).

Moreover, Bolton et al. (Bolton, Smith, and Jacobson 2004) highlight how public procurement policies, such as set-asides and subcontracting mandates, have been used to improve small business access to government contracts. These insights further reinforce the need for targeted reforms in procurement systems to facilitate greater participation from small firms. The UK's experience with supplier diversity initiatives is particularly relevant, as evidenced by Bolton et al. (Bolton, Smith, and Jacobson 2004), who emphasized how diversity programs and government support have led to better outcomes for smaller suppliers.

The results of this study support these conclusions, highlighting that high-value contracts, particularly in specific categories, still present significant challenges for small businesses. There is also a strong case for implementing additional policy measures to ensure fairer representation in the procurement process.

By building on these findings, this paper underscores the importance of continuous policy interventions to mitigate the structural barriers small businesses face in public procurement.

## A.6 Future Directions: Enhancing Survey and Data Collection Methods

While this study provides valuable insights into the disparities in procurement opportunities for small businesses, future research could build on this work by incorporating additional survey data and experimental designs. For instance, surveys could be conducted with procurement officers to better understand the barriers they perceive in awarding contracts to small businesses. Furthermore, experimental studies could be used to test the effectiveness of specific policy interventions, such as training programs or set-aside policies, in improving small business participation.

## B Model details

### B.1 Variables

These variables that was not used in the analysis, but are essential for understanding supplier representation in the City's public procurement process. These variables include:

- **Unique Row Identifier (id):** A unique identifier for each record in the Open Data database.
- **Composite Key (unique\_id):** A unique identifier combining related procurement details.
- **Document Number:** A reference number used by the City to track the solicitation during the procurement process.
- **Division:** The City Division requesting the procurement. The awarded supplier will have a contract with this division.
- **Buyer Details:** Includes the buyer's name, email, and phone number as the primary purchasing contact for the solicitation or contract.
- **Solicitation Document Description:** A summary of the solicitation, providing additional context about the procurement request.

### B.2 Details on Selection of Big Business Out From Data

The reult of selection of big business is by using ChatGPT.

I asked: in the data, under the Successful\_Supplier, can you anaysis and pull out the excat names that are consider as big business and put them into a big\_business\_list, so i can use it for r code

ChatGPT said: ChatGPT The dataset contains the following columns:

unique\_id: A unique identifier for each record. Document\_Number: The associated document number. RFx\_Type: Type of request (e.g., RFQ, RFP). High\_Level\_Category: The category of the procurement. Successful\_Supplier: The name of the supplier who won the contract. Awarded\_Amount: The monetary value of the awarded contract. Award\_Date: The date the contract was awarded. Division: The division responsible for the contract. Small\_Business: Indicates if the supplier is classified as a small business (True/False). To identify big businesses, I will filter out suppliers where Small\_Business is False and create a list of their names.

The following are the big businesses identified from the dataset:

Guillevin International Co. Sysco Sysco Toronto, A Division Of Sysco Canada Inc. Stantec Consulting Ltd KPMG LLP KPMG Bennett Mechanical Installations Fer-Pal Construction Ltd. WSP Canada Inc Black And McDonald Limited Thermo Fisher Scientific Sutherland-Schultz Ltd. Morrison Hershfield Ltd Morrison Hershfield Metro Freightliner Hamilton Inc Ricoh Canada Inc Altus Group Limited Kroll Consulting Canada Co. Suncorp Valuations Logixx Security Inc Garda Canada Security Corporation Drake International Inc Bevertec CST Inc AECOM Canada Ltd Stericycle Ulc Schindler Elevator Corporation Graham Bros. Construction Limited Morrison Hershfield Limited Softchoice LP Damen Shipbuilding 5 B.V. Graham Bros Construction Ltd. Gartner Canada Parsons Inc Jacobs Consultancy Canada Inc Stericycle, Ulc

This list is now ready for use in your R code. Let me know if you need it in a specific format!

## B.3 Diagnostics

### Check Model Fit Using AIC

The AIC (Akaike Information Criterion) is a measure of model quality, where lower values indicate a better fit, while penalizing for model complexity. To assess model performance, we compare the AIC values of different models, including those with additional predictors or terms. A lower AIC suggests a better fit. In our comparison, we evaluated the base model alongside two other models that include additional terms shown in Table 6. The base model, with the lowest AIC, demonstrates the best performance and is therefore used as the final model in this paper.

Table 6: AIC Comparison of Logistic Regression Models

Model Description	AIC
Model 1: Base Model: RFx_Type + High_Level_Category + Awarded_Amount	254.67
Model 2: With Division	284.49
Model 3: With Interaction Term: RFx_Type * Awarded_Amount	257.65

### Hosmer-Lemeshow Test

The Hosmer-Lemeshow test evaluates whether the observed outcomes align with the predicted probabilities, making it particularly useful for logistic regression models. A p-value greater than 0.05 suggests that the model fits the data well, as the observed and predicted values are closely aligned. In our analysis (Table 7), the Base model yielded a p-value of 0.7, indicating a good fit to the data.

Table 7: Hosmer-Lemeshow Goodness-of-Fit Test Results

Metric		Value
X-squared	Chi-squared	5.4906
df	Degrees of Freedom	8.0000
	P-value	0.7041

### Residual Deviance

Residual deviance compares the deviance of the model to that of the null model (a model with no predictors). A lower residual deviance indicates a better fit. A p-value less than 0.05 suggests that the model significantly improves upon the null model. As shown in Table 8, the model significantly outperforms the null model, with a p-value of zero.

Table 8: Deviance-Based Model Comparison

Metric	Value
Null Deviance	288.5661
Residual Deviance	236.6693
Degrees of Freedom	8.0000
P-Value	0.0000

## References

- Bolton, S., L. Smith, and P. Jacobson. 2004. "Policy Approaches to Supplier Diversity in Public Procurement." *Public Administration Review* 64 (6): 824–32.
- OpenAI. 2023. "ChatGPT: A Language Model." <https://www.openai.com/chatgpt>.
- Robinson, David, Alex Hayes, Simon Couch, and Julia Silge. 2023. *Broom: Convert Statistical Objects into Tidy Tibbles*. <https://CRAN.R-project.org/package=broom>.
- Toronto, City of. 2024. "TOBids Awarded Contracts." <https://open.toronto.ca/dataset/tobids-awarded-contracts/>.
- Wickham, Hadley et al. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Wickham, Hadley. 2023. *Ggplot2: Elegant Graphics for Data Analysis*. <https://ggplot2.tidyverse.org/>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, and Winston Chang. 2023. *Scales: Scale Functions for Visualization*. <https://CRAN.R-project.org/package=scales>.
- Wickham, Hadley, and Garrett Golemund. 2023. *Lubridate: Make Dealing with Dates a Little Easier*. <https://CRAN.R-project.org/package=lubridate>.
- Witt, M., L. Smith, and P. Johnson. 2019. "Barriers to Small Business Participation in Government Procurement." *Journal of Public Procurement Policy* 11 (2): 125–42.
- Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.
- Zhu, Hao. 2023. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.