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

Mandy He

December 3, 2024

This study explores the representation of small businesses in Toronto's public procurement system, specifically examining how high-value contracts—particularly in construction services—tend to favor larger firms. Using data from the Toronto Bids Awarded Contracts dataset, the analysis uncovers significant disparities in contract values and supplier participation, revealing that small businesses face considerable barriers in accessing these high-value opportunities. The findings highlight structural inequities that limit small business inclusion and point to the need for reform in procurement policies. To create a more equitable procurement environment, the study recommends enhancing transparency, improving outreach, and providing stronger support for small businesses.

Table of contents

---

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

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 substantial resources to suppliers for goods and services, providing essential economic opportunities. For small businesses and historically underrepresented sectors, public contracts represent vital opportunities for growth and success. However, ensuring 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 contract values, 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, offering 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 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 such as construction services, which 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, as well as 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 help reduce economic inequality. The study highlights 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 and the model, 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, including 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, several packages in the R programming language (R Core Team 2023) 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, the language model LLM - ChatGPT 4o mini (OpenAI 2023) - was utilized to assist with coding tasks, including data cleaning, simulating missing values, generating graphs, and refining text.

Table 1 provides a preview of the clean data that will be used throughout the analysis in 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 essential 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 with identifying procurement needs, issuing requests for proposals (RFPs), evaluating bids, and ultimately awarding contracts to successful suppliers. Each stage reflects real-world events that are systematically documented and incorporated into the dataset.

The transition from real-world activity to structured data begins when the City identifies a need, such as for infrastructure development, service improvements, or administrative support. The procurement process starts with the issuance of a solicitation document that outlines the scope of work, evaluation criteria, and specific project or service requirements. Suppliers then submit their bids, which are evaluated based on predefined criteria, such as cost, quality, and compliance with 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 awarded contract amount, the supplier's name, and the City division requesting the procurement—are recorded for documentation and analysis.

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

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

Ultimately, the dataset serves as a structured representation of the City's complex procurement process. Each entry 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 note that, while detailed, the 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 essential for understanding supplier representation in the City's public procurement process. The primary variables 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 variable indicates the process used to seek bids or proposals from suppliers.

- **High-Level Category:** This variable groups the goods or services being procured, such as Goods and Services, Professional Services, or Construction Services.

- **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 enables 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. All other variables that the dataset compile can be found in Section B.1.

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 valuable 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 provide 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 achieve this, a logistic regression model is used, as it is well-suited 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:

$$\log\left(\frac{P(\text{Small\_Business} = 1)}{1 - P(\text{Small\_Business} = 1)}\right) =$$

$$\beta_0 + \beta_1 \cdot \text{RFx\_Type} + \beta_2 \cdot \text{High\_Level\_Category} + \beta_3 \cdot \text{Awarded\_Amount} + \beta_4 \cdot \text{Award\_Date}$$

This equation uses the log-odds transformation (logit function) to model the probability of a contract being awarded 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).

β_0**:** The intercept of the model.

β_1**:** The coefficient for RFx_Type, representing the effect of the solicitation type on the likelihood of awarding a contract to a small business.

β_2**:** The coefficient for High_Level_Category, representing the effect of the procurement category (e.g., construction, professional services) on small business participation.

β_3**:** The coefficient for Awarded_Amount, representing how the contract value influences the probability of a small business being awarded the contract.

β_4**:** The coefficient for Award_Date, representing how time influences the probability of a small business being awarded the contract.

The interpretation of coefficients:

β_1**:** Change in log-odds of awarding a contract to a small business for different RFx types.

β_2**:** Change in log-odds for high-level categories.

β_3**:** Effect of the awarded amount on the log-odds of awarding a contract to a small business.

β_4**:** Change in log-odds associated with each unit increase in the date

2.5 Model Setup

The outcome variable is modeled using a Bernoulli distribution:

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

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

$$\mu_i = \text{logit}^{-1}(\alpha + \beta_1 \cdot \text{RFx\_Type}_i + \beta_2 \cdot \text{High\_Level\_Category}_i + \beta_3 \cdot \text{Awarded\_Amount}_i) + \beta_4 \cdot \text{Award\_Date}_i)$$

This ensures that μ_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, making it well-suited 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 evaluated using standard techniques such as AIC, the Hosmer-Lemeshow test, and residual deviance. Diagnostics for these tests are presented in Section B.3.

Additionally, we assume that model assumptions, such as the linearity of the log-odds and the independence of observations, hold.

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 considered was 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 limited 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 not available 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 contracts awarded, suggesting meaningful representation in Toronto's procurement process. This observation points to potentially equitable opportunities for smaller firms, though further analysis is needed to confirm the systemic accessibility of these opportunities.
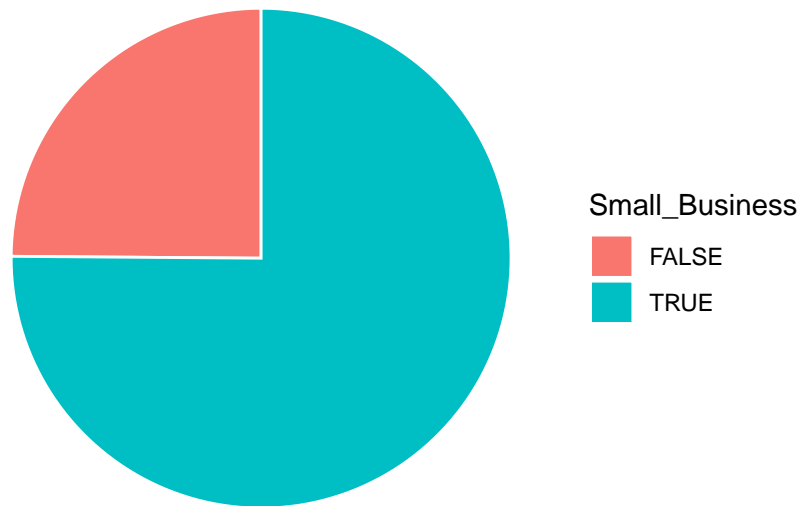
Proportion of Small Business Awards

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 total awarded amounts across various RFx types, revealing several important trends in the procurement process. First, the Request for Tender (RFT) stands out as the dominant procurement method, with a total awarded amount exceeding $1 billion. This suggests that tenders are the primary mechanism for large-scale city contracts. Following RFT, the Request for Quotation (RFQ) and Request for Proposal (RFP) types each account for awarded amounts just below $375 million, indicating their significant but 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 minimal awarded amounts. This suggests that these methods are infrequently used in awarding contracts. Overall, the chart highlights a clear preference for tenders and quotations in the city's procurement activities, with a focus on larger, more formal procurement processes.
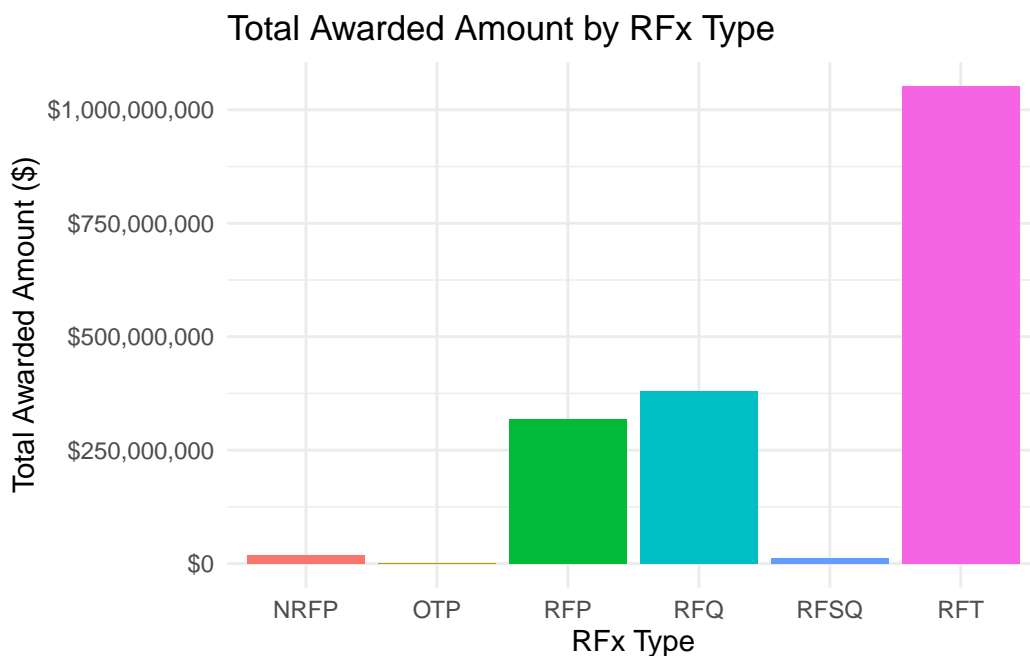

Figure 2: Total awarded amounts across RFx types show that Request for Tender (RFT) contracts dominate the procurement budget, with over $1 billion in awarded funds. Request for Quotation (RFQ) and Request for Proposal (RFP) contracts follow, each with awarded amounts just below $375 million. Other RFx types, including NRFP, OTP, and RFSQ, receive minimal awarded amounts, indicating that 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 the RFQ, RFT, RFP, and RFSQ categories, consistently securing the majority of contracts. In contrast, NRFP awards are almost exclusively given to Large Businesses, while OTP contracts are entirely awarded to Small Businesses.

Proportion of Small vs. Large Businesses by RFx Type
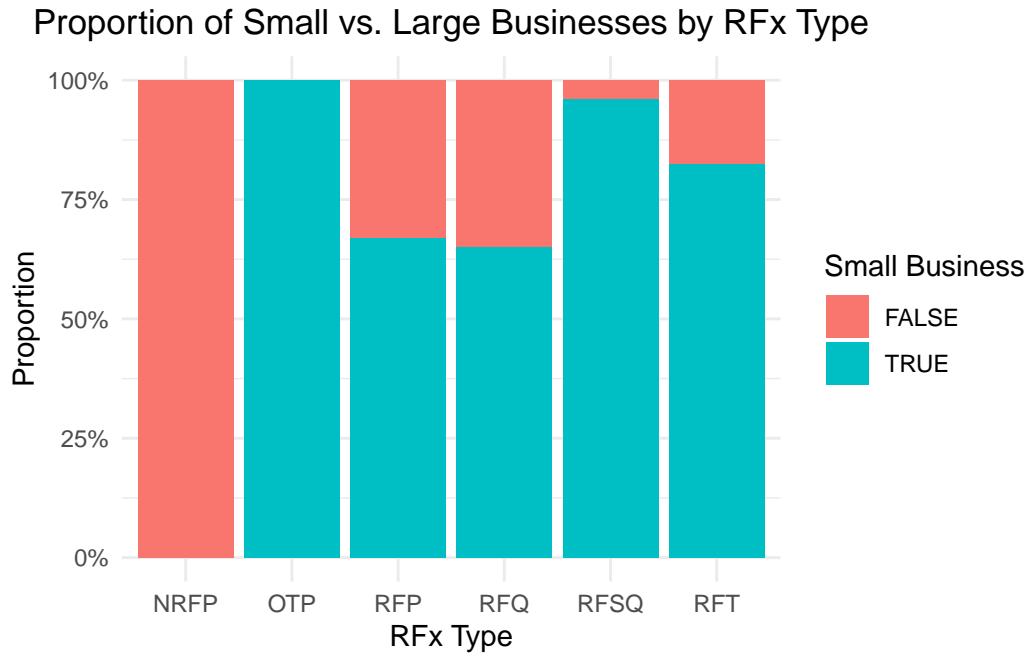
Figure 3: Supplier classification (Small Business vs. Large Business) across different RFx types. Small businesses (blue) dominate RFQ, RFP, and RFSQ contracts, reflecting higher participation in these categories. In contrast, Large businesses (red) are predominantly awarded NRFP and OTP contracts, with 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 are moderately represented in RFQ and RFP, but their awarded amounts remain significantly lower. These disparities suggest that factors such as contract scale, complexity, or financial prerequisites may favor large businesses. The data points to potential systemic barriers preventing small businesses from accessing higher-value contracts.
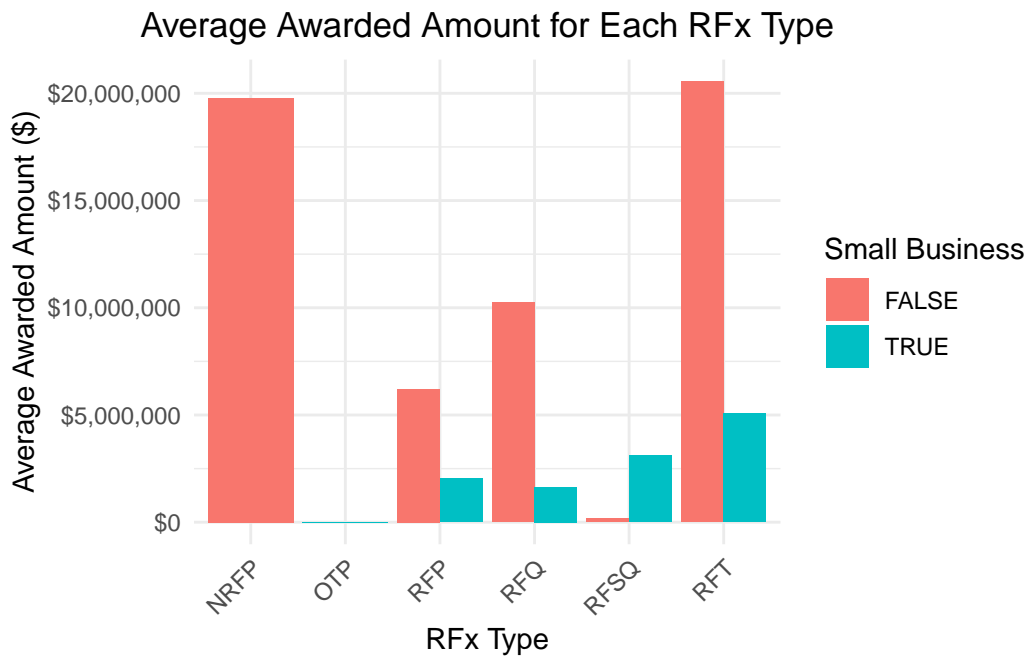

Figure 4: Awarded amounts to small and large businesses across various RFx types. Large businesses are awarded significantly higher contract values, particularly in RFT and NRFP categories, where high-value contracts are more common. Small businesses are comparatively underrepresented in these categories 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 exceeding $75 million. This suggests 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, however, shows minimal representation, with awarded amounts clustered at lower values, indicating that it is infrequently used for high-value procurement.
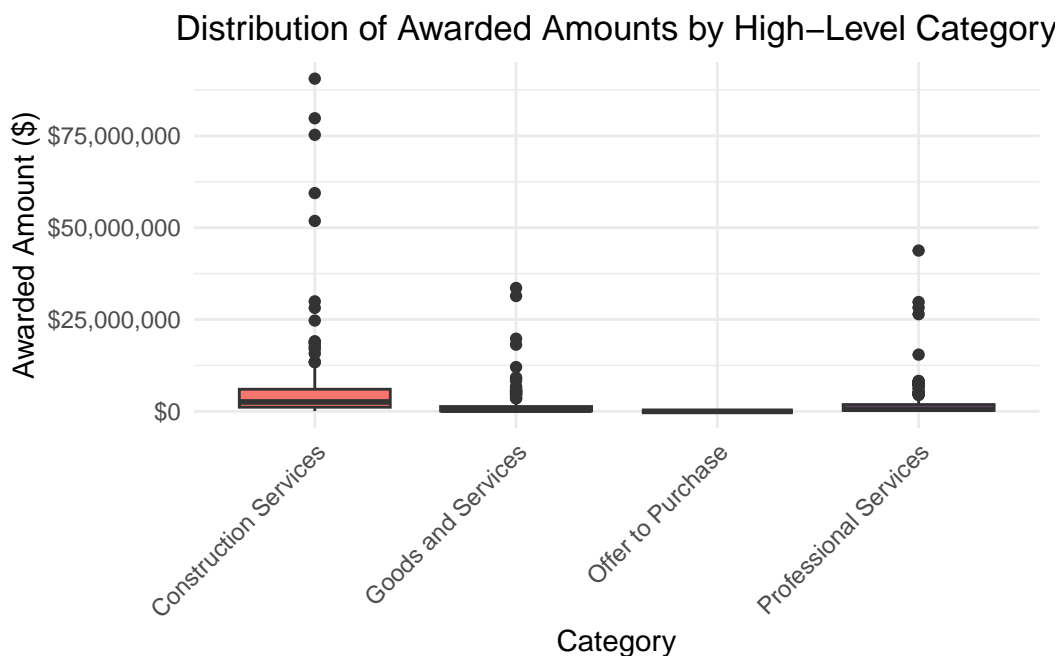
Figure 5: Distribution of awarded contract amounts across high-level procurement categories such as Construction Services, Goods & Services, Offer to Purchase, and Professional Services. Construction Services show the widest range, with some contracts exceeding $75 million, while other categories, such 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 both the Construction Services and Professional Services categories, large businesses dominate, suggesting that small businesses may face barriers to participation, such as stringent qualification requirements or the scale of projects that favor larger firms. In contrast, the Goods & Services and Offer to Purchase categories show a higher proportion of contracts awarded to small businesses, indicating that these categories are more inclusive, with fewer financial or operational barriers, making them more accessible to smaller enterprises. Overall, these trends highlight varying levels of inclusivity across procurement sectors, with some categories offering more opportunities for small businesses to engage in the process.
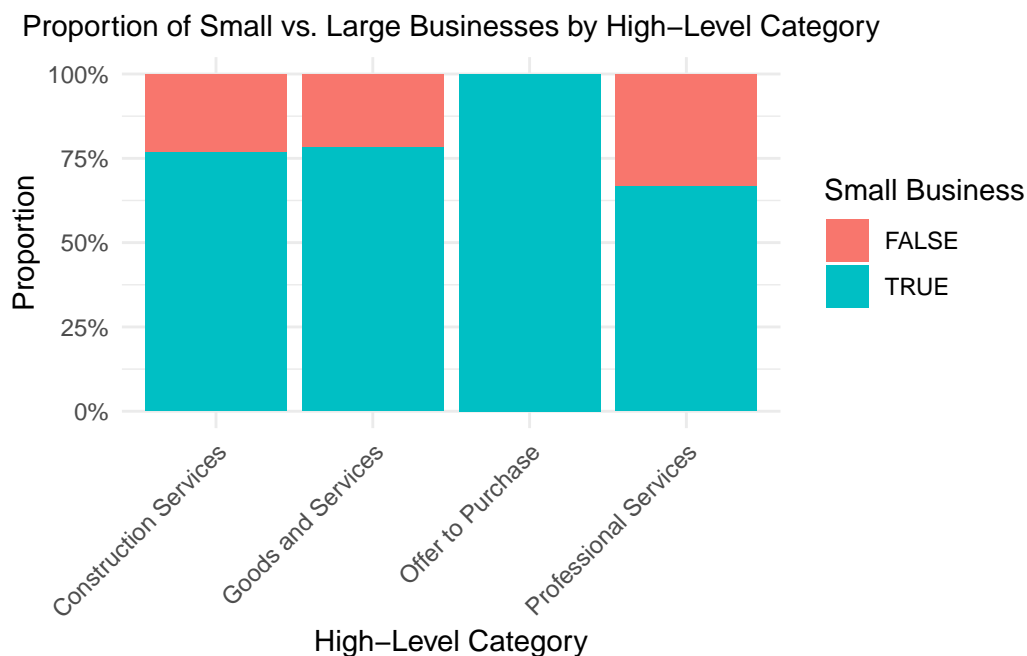
Figure 6: Proportion of contracts awarded to small businesses (in blue) versus large businesses (in red) across procurement categories. Construction Services and Professional Services show a larger share of contracts awarded to large businesses, indicating potential barriers for small businesses in these 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 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 disparity suggests that large-scale contracts in this category likely require resources or qualifications 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 fewer barriers for small firms and greater competition. The Offer to Purchase category shows no recorded average awarded amounts, suggesting minimal or no activity within this category in the dataset. In Professional Services, large businesses still receive higher average awards, though the gap is smaller compared to Construction Services, indicating that while small businesses have some opportunities, disparities in awarded amounts persist. These findings highlight significant variations in contract accessibility for small businesses across procurement categories, with some offering more inclusive opportunities than others.
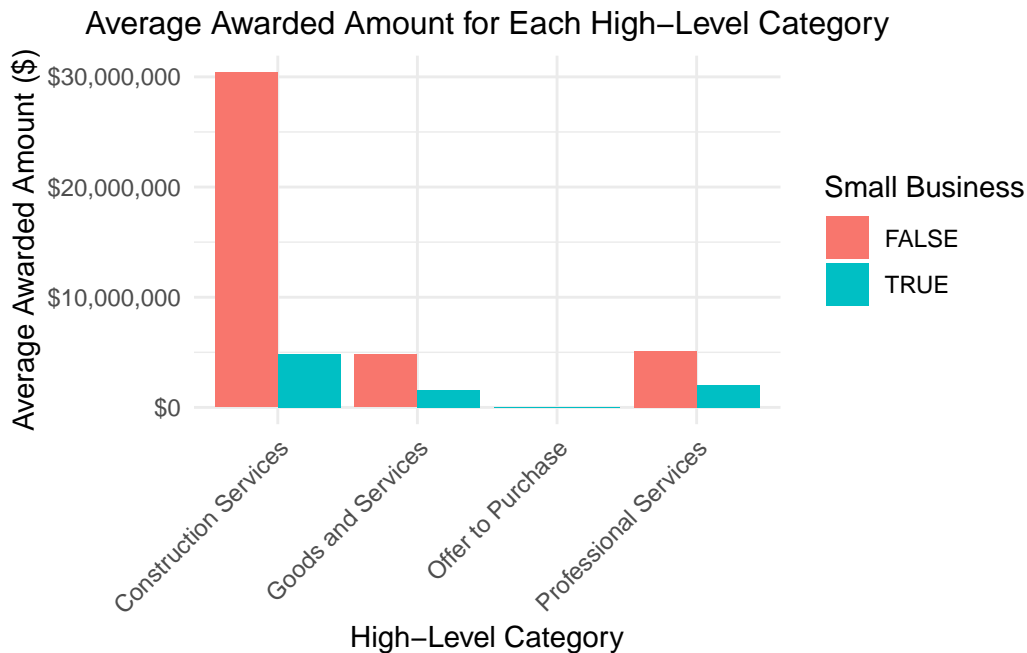
Figure 7: Average awarded amounts for small and large businesses across various contract categories. Large businesses dominate the Construction Services category with average awards nearing $30 million, while small businesses receive considerably lower awards. Goods & Services contracts show more balanced awards, offering similar average values for both small and large businesses, while Professional Services shows a smaller gap favoring large businesses.

17

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 contribute to 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 slight increases in late 2023 and mid-2024, suggesting modest improvements in awarded amounts for small businesses during these periods. The overall trends highlight a consistent disparity between small and large businesses, with large businesses securing the majority of high-value contracts. However, the relatively stable awarded amounts for small businesses indicate that, while their opportunities remain limited, there has been a steady, if modest, flow of contracts to smaller enterprises.
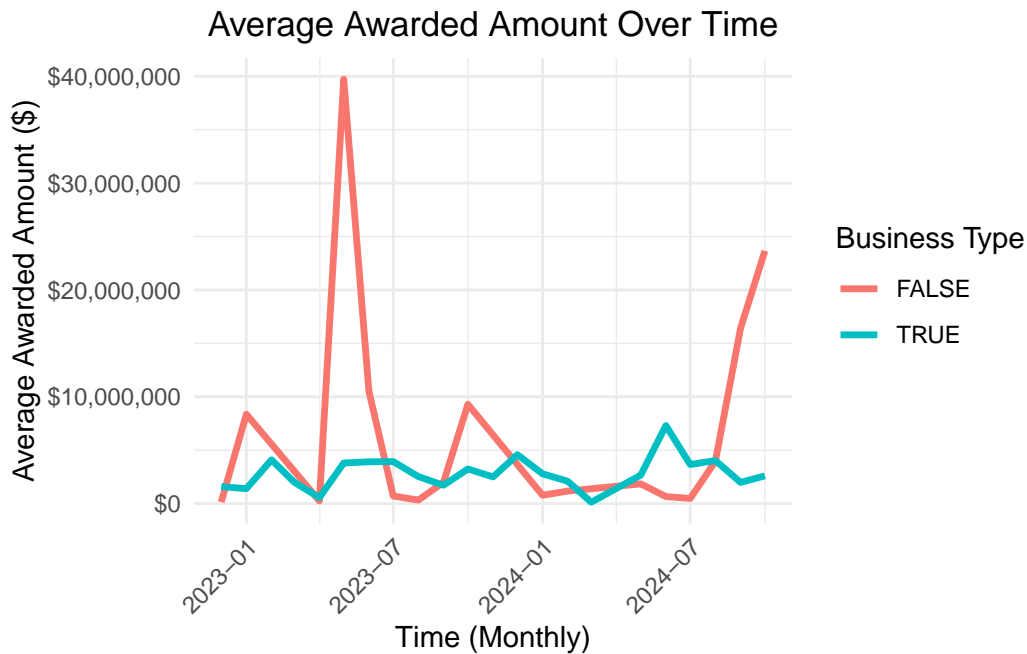

Figure 8: Average awarded amounts to small and large businesses over time. Large businesses experience notable fluctuations, with significant peaks in mid-2023 and late 2024, reflecting the occasional awarding of high-value contracts. Small businesses show a steady trend of lower awarded amounts, mostly 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 period. Most awarded amounts are concentrated below $25 million, suggesting that high-value contracts are relatively rare. Among the few high-value contracts, large businesses dominate, receiving a disproportionate share of the larger awards. Small businesses, while represented across the timeline, are consistently awarded amounts within lower ranges, 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.
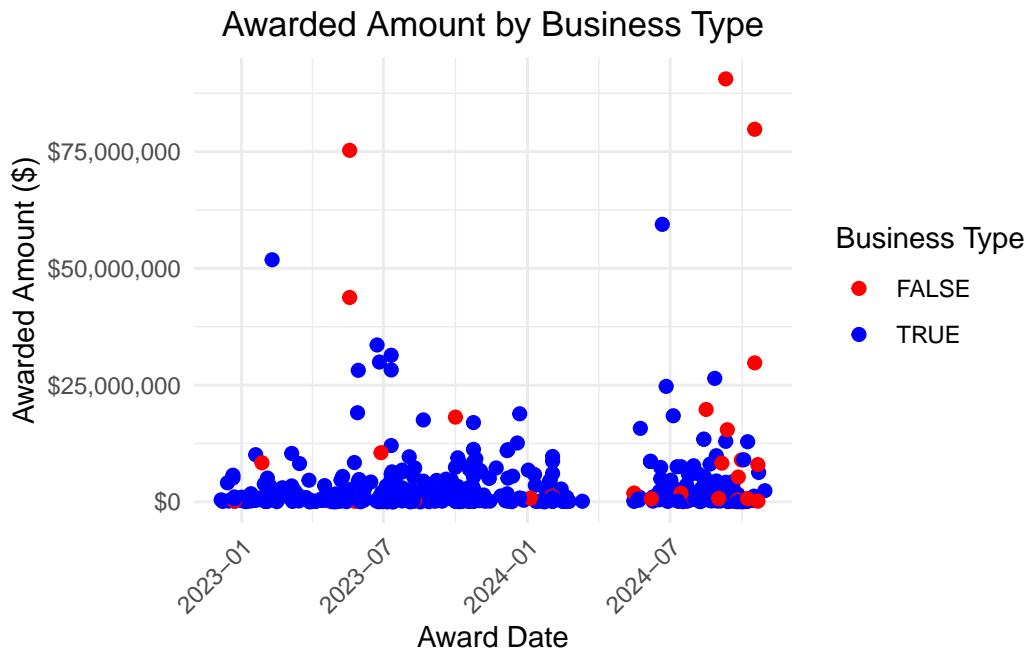

Figure 9: Scatter plot showing individual awarded amounts to small and large businesses over time. The majority of awarded amounts remain below $25 million, with only a few businesses receiving higher amounts. Large businesses tend to secure higher awards, as seen by their concentration at the upper 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 shows a positive trend, indicating a gradual increase in awarded amounts to small businesses, which signals some progress in their procurement outcomes over the analyzed period. However, the spread of data points around the regression line reveals considerable variability in monthly awarded amounts, with certain months exhibiting 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. These findings emphasize the need for consistent policy measures to support and sustain equitable participation by small businesses in procurement activities.

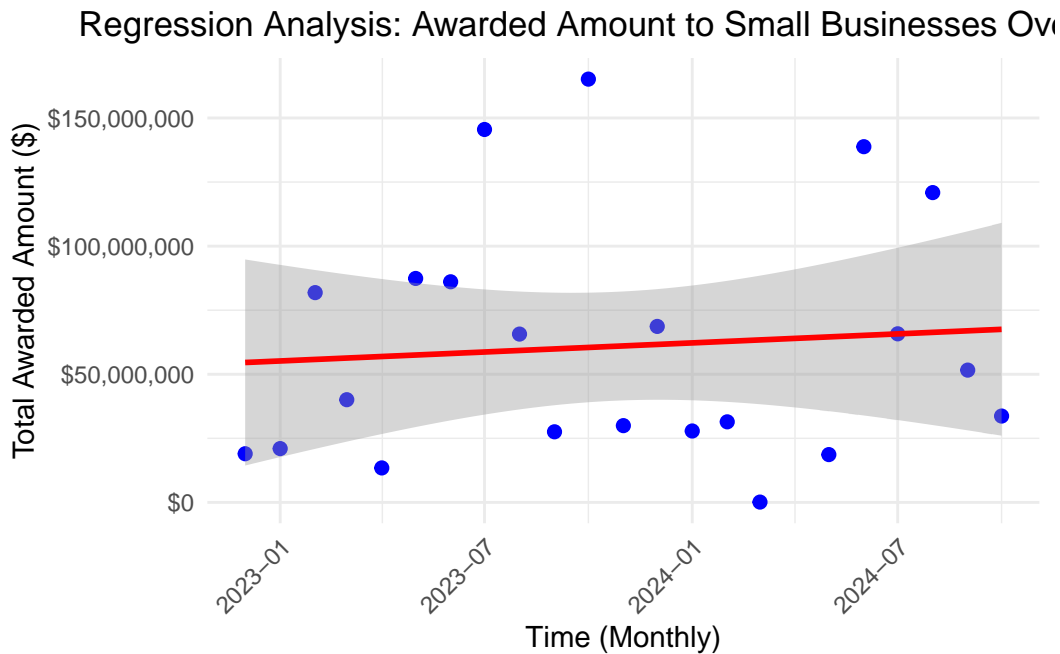Figure 10: Trend of awarded amounts to small businesses over time, as shown by the regression line. The gradual upward trend suggests a slight increase in awarded amounts to small businesses, though with considerable variability in the data points, indicating inconsistent opportunities. This variability 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)	75.9299	1455.5273	9.460847e+32	0.0000	Inf	0.9584
RFx_TypeOTP	25.4890	2058.2438	1.174217e+11	0.0000	Inf	0.9901
RFx_TypeRFP	15.8649	1455.3980	7.763294e+06	0.0000	Inf	0.9913
RFx_TypeRFQ	15.6943	1455.3976	6.545610e+06	0.0000	Inf	0.9914
RFx_TypeRFSQ	11.8139	1455.3993	1.351119e+05	0.0000	Inf	0.9935
RFx_TypeRFT	12.7427	1455.3987	3.420541e+05	0.0000	Inf	0.9930
High_Level_CategoryGoods and Services	-4.4131	1.7801	1.210000e-02	0.0004	0.3969	0.0132
High_Level_CategoryOffer to Purchase	NA	NA	NA	NA	NA	NA
High_Level_CategoryProfessional Services	-5.0685	1.7306	6.300000e-03	0.0002	0.1870	0.0034
Awarded_Amount	0.0000	0.0000	1.000000e+00	1.0000	1.0000	0.0000
Award_Date	-0.0043	0.0010	9.957000e-01	0.9938	0.9976	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, provide several key insights:

Impact of Awarded Amount: This is statistically significant. A unit increase in the awarded amount corresponds to no change in the odds ratio (odds ratio = 1.0000), which suggests that the awarded amount does not influence the likelihood of a supplier being a small business.

High-Level Category Influence: Certain procurement categories, such as "Goods and Services" and "Professional Services," significantly reduce the probability of small businesses winning contracts. This points to an uneven playing field in these sectors, likely driven by higher technical, operational, or financial requirements that make it more difficult 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 than the specific procurement method used.

Impact of Award Date: This is statistically significant. A one-unit increase in Award_Date (e.g., one day) slightly decreases the odds of a supplier being a small business (odds ratio = 0.9957, approximately a 0.43% decrease per day). Over time, the likelihood of small businesses winning contracts appears to decrease slightly, indicating potential structural trends or barriers developing in procurement.

Implications: The findings highlight the substantial barriers small businesses face, particularly in securing high-value contracts and in procurement categories that require 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.

4 Discussion

4.1 Scope and Approach of the Study

This study examines the representation of small businesses in Toronto's public procurement system across various RFx types (Request for X) and high-level procurement categories. It combines 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 instance, the bar charts (Figure 3, Figure 4) illustrate how RFx types like RFT disproportionately favor large businesses, while RFQ and RFP categories provide more balanced opportunities. However, the logistic regression model (Table 5) does not show that RFx types significantly impact small business success. The analysis suggests that the structure of RFx types alone does not account for the likelihood of small businesses winning contracts. Larger awarded amounts and more complex procurement categories still seem to reduce the likelihood of small businesses securing contracts, although Awarded Amount itself was not statistically significant in the model.

4.2 Insights on Small Business Representation in Procurement

The analysis reveals that small businesses achieve the greatest success in procurement categories and RFx types with less complex requirements, such as RFQ (Request for Quotation) and RFP (Request for Proposal). As shown in the stacked bar chart (Figure 3), these RFx types lower barriers to entry by focusing on straightforward pricing and proposal quality, enabling smaller firms to compete effectively. In contrast, RFT (Request for Tender) contracts, which often involve high-value, resource-intensive projects, are dominated by large businesses due to their ability to meet stringent requirements and offer economies of scale. While the logistic regression results do not show that RFx type significantly impacts small business participation, the descriptive analysis still supports the idea that RFQ and RFP categories may offer more opportunities for small businesses.

The regression model found that the Awarded Amount itself does not significantly affect the likelihood of small business success, contradicting the earlier suggestion that larger contract values create barriers. However, procurement categories like Professional Services continue to show challenges for small businesses, confirming the earlier observation that more complex categories tend to favor larger firms.

These findings directly address the research question by identifying how the structure of RFx types and procurement categories impacts small business participation in public procurement. The observed disparities highlight systemic inequities that hinder small businesses from competing in high-value contracts, reinforcing the need to evaluate whether public procurement policies are genuinely inclusive. This analysis provides evidence to support the argument that procurement reforms, such as simplifying tender requirements or creating small business-specific RFx categories, could foster greater equity and inclusivity in the system.

4.3 Uneven Trends Over Time in Small Business Success

The research also reveals that while small businesses show some progress in securing contracts over time, the overall trend remains uneven. A slight increase in awarded amounts to small businesses is visible in the line graph (Figure 8), indicating gradual growth, but volatility in contract values persists. The scatter plot with a regression line (Figure 10) further highlights the inconsistency in awarded amounts over time. These findings suggest that, despite some improvements, opportunities for small businesses remain unstable and inconsistent, influenced by factors such as RFx type and procurement category. The regression model also found that Award Date has a significant negative effect on small businesses, indicating that over time, small businesses are less likely to be awarded contracts.

These findings relate directly to the research question by emphasizing how structural factors—such as RFx type and procurement category—continue to shape small business outcomes unpredictably. For example, while certain policies may temporarily boost access, systemic issues like complex eligibility requirements or lack of targeted programs limit sustainable growth. This underscores the need for consistent policy measures designed to stabilize and enhance small business participation, aligning with the overarching goal of fostering equitable access in public procurement.

4.4 Weaknesses and Next Steps

Despite its comprehensive approach, the study has several limitations that highlight areas for future exploration. A significant shortcoming is the lack of qualitative data, which could offer deeper insights into the barriers small businesses face in the procurement process. For instance, while the box plot (Figure 5) illustrates variations in awarded amounts across procurement categories, it cannot address subjective factors like bidders' experience or the perceived complexity of tender processes. Similarly, while the logistic regression model (Table 5) identifies key statistical relationships, it cannot account for external influences, such as shifts in government policy or broader economic conditions, that impact small business participation.

To address these gaps, future research should prioritize qualitative studies, such as interviews or case studies, to uncover nuanced challenges like eligibility requirements, administrative burdens, or document complexity. Expanding the analysis to multiple cities or regions could provide comparative insights into the effectiveness of procurement policies and their impact on small businesses. Longitudinal studies tracking policy changes over time would also be valuable in assessing their success in improving small business access. For example, the time series analysis (Figure 8) reveals fluctuations in awarded amounts, underscoring the need for investigations into policy interventions and their sustained effects. Additionally, studying government initiatives aimed at supporting small businesses can help identify strategies that effectively mitigate systemic inequities and foster equitable access in procurement.

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 the representation of small businesses in Toronto's public procurement system. The dataset consists of contracts awarded across various RFx types, categorized by business size and procurement category. By focusing on small versus large business participation, we highlight disparities that may point to systemic barriers faced by small businesses. Understanding the methodologies used to gather and analyze this data is essential to ensuring the robustness of our findings.

The data used in this study is observational, collected from publicly available procurement records over a defined period. Unlike survey data, which involves actively soliciting responses from participants, observational data in procurement enables us to examine real-world behaviors—specifically, contract awards—as they occur naturally, without any intervention.

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

While the procurement data we analyze is observational, understanding survey sampling principles can enhance our interpretation of the results. Procurement records may be viewed as a form of "sample" from the broader population of potential contracts. By applying proper sampling techniques, we can ensure that our analysis more 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), procurement category (e.g., construction services, professional services), or business size (small vs. large). Stratified sampling ensures that each category is adequately represented, improving the accuracy of assessments of disparities in contract awards.

For instance, if small businesses are predominantly represented in RFQ contracts but underrepresented in RFT contracts, stratified analysis would allow for a more precise comparison across procurement categories, helping to uncover systemic barriers that may be affecting small business participation.

Random Sampling vs. Observational Data

While random sampling is a key technique in survey design 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 intentional sampling. However, caution is needed to account for other forms of bias that could affect the data collection process, such as data reporting or timing biases.

Unlike random sampling, observational data may not capture all segments of the population (e.g., missed or incomplete contract records). Thus, the limitations of the data should be carefully considered when drawing conclusions. If additional survey methods were used—such as surveying procurement officials or small business owners about their experiences—this could complement the observational data and provide further insights into potential gaps.

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

Measurement Issues

Procurement data relies on contract records that include values, RFx types, and business classifications. However, measurement errors can occur due to inconsistent reporting, differences in how contracts are categorized, or incomplete data entries. In this study, we validated the data to ensure it is accurate and up-to-date. However, it is important to recognize potential measurement issues, particularly when interpreting trends or 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 how small business participation is assessed across different procurement categories. Moreover, some contracts may involve multiple suppliers, which could lead to variations in how contract values are reported.

Data-Related Biases

Although observational data avoids some biases inherent in survey methodologies, it is not immune to selection biases. In procurement, certain contracts may be more likely to be awarded to large businesses due to the scale or complexity of the work (e.g., large infrastructure projects), which could distort our understanding of small business participation. Additionally, small businesses may face greater barriers to participating in high-value contracts, which could further skew the data on their representation.

To address these potential biases, our study employs regression techniques to control for variables like awarded amount and procurement category. This helps isolate the effect of business size on contract awards. However, further refinement of the analysis could include techniques like propensity score matching or other methods to correct for potential selection biases and ensure a more accurate representation of small business participation.

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) might 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 reductions in administrative burden), 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 such as 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, showing that high-value contracts, particularly in certain categories, continue to 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 that 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 these findings by incorporating additional survey data and experimental designs. For instance, surveys could be conducted with procurement officers to gain a deeper understanding of the barriers they perceive when 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.

- **Successful Supplier:** The name of the supplier awarded the contract.

- **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 Steps Taken to Classify Small and Large Businesses

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

The classification of big businesses from the raw data was based on a combination of factors, including prior knowledge of well-known large firms, research into supplier names, and

thresholds such as contract size and industry dominance. Suppliers recognized as multinational, high-revenue, or dominant in sectors like construction and consulting were flagged as big businesses. Additional verification through public records and analysis of awarded amounts helped refine the classification, ensuring it accurately reflected the scale and market presence of these firms. This process created a foundation for distinguishing big businesses from smaller enterprises in the dataset.

Steps:

1. Manual Selection Based on Known Characteristics:

Using prior knowledge, research, or publicly available information, suppliers were classified as big businesses if they were recognized as large, established firms. These companies might be well-known for their market presence, large-scale operations, or participation in major contracts (e.g., KPMG, AECOM, Stantec). Indicators such as the company being multinational, having high revenues, or being dominant in certain industries (e.g., construction, consulting) informed their classification.

2. Threshold-Based Classification (If Applicable):

For some businesses, thresholds like contract size, revenue, or number of employees might have been inferred or estimated. Suppliers with consistently high-value contracts (e.g., those often exceeding $1 million) were assumed to be large businesses.

3. Categorization Based on Supplier Type:

Specific industries or sectors (e.g., large construction firms, multinational consulting companies) were likely flagged as dominated by big businesses. For example: Construction firms: Companies like Graham Bros. and Black & McDonald Limited were classified as big due to their role in large-scale projects. Technology and consulting firms: Firms like Softchoice, KPMG, and Jacobs Consultancy were marked as big businesses due to their industry prominence.

4. Review of Contract Values:

Suppliers awarded contracts with very high total values (e.g., in the millions) were flagged as large businesses. This assumption ties contract value to company size. Public Records and Research:

For any ambiguous suppliers, online searches or business directories (e.g., company websites, municipal reports) might have been used to verify the scale of the business. Companies with a national or international footprint or high revenue streams were classified as large.

5. Iteration and Refinement:

After creating an initial classification of big businesses, further refinement could involve cross-checking against supplier names, award amounts, or external benchmarks.

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-fitting model. In our analysis, we compared the base model with two other models that included additional terms, as shown in Table 6. The base model, with the lowest AIC, demonstrates the best fit and is therefore selected as the final model for this study.

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 how well the observed outcomes align with the predicted probabilities, making it especially useful for assessing the fit of logistic regression models. A p-value greater than 0.05 suggests that the model fits the data well, indicating that the observed and predicted values are closely aligned. In our analysis (Table 7), the base model yielded a p-value of 0.7, which suggests 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.

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

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 Grolemund, 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 Grolemund. 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.