Spring Bank Drive Report

Group 7

November 27, 2022

1. Background and Motivation

The Canada-Ontario Infrastructure Program planned to widen 2.55 kilometers of Springbank Drive from two lanes to four between Horton Street and Wonderland Road and to construct turning lanes at critical intersections. The construction is expected to be completed by 2004.

The main issue from this construction was the compensation the residents could seek according to the Expropriations Act for the expected front yard setbacks, increased traffic volumes, and other problems caused by the widening of Springbank Drive due to the land expropriation already approved by the city to allow construction to proceed. In the past, the city used a formula based on the market value of the amount of land expropriated plus a lump sum that included costs for all loss, disturbance costs, and injurious affection. However, we did not feel that the city's formula properly accounted for all the factors and the total loss of market value the residents would suffer, leaving the city vulnerable to compensation claims.

From the history of sales price data around Springbank Drive, we will estimate the loss of market value for each homeowner and justify the approach in case the City of London is challenged over the compensation claims in court.

2. Descriptive Statistics

Table 1: Descriptive Statistics

Variables	Mean	Standard Deviation
Average.View	0.58	0.50
Good.View	0.22	0.42
Average.Interior.Condition	0.45	0.50
Good.Interior.Condition	0.34	0.48
Excellent.Interior.Condition	0.12	0.32
LANESRD (Four Lane Road)	0.33	0.47
LFA (Lot Frontage Area)	902.84	255.55
BSMTFINAREA (Basement Area)	396.78	287.16
TRAFCOUNT (Traffic Count)	20221.15	8368.87
PRICE (Sales Price)	134469.20	27743.65

The sample has 104 observations of the sales price of residential properties on and around Springbank Drive from January 1998-May 2003. The data included past sales of residential properties. A total of 8 attributes were included for each property and the sales price. The data is taken from the Company Files.

Approximately 58% of the houses have an average view, 22% have a good view, and the remaining have a fair view. 45% of houses are in average interior condition, 34% have a good condition, 12% are in excellent condition, and the remaining homes are in fair condition. 33% of the houses have access to a four-lane side road, the average frontage area of the house is 902.84 square feet and the max frontage area is around 1,674 square feet. The average basement area of the houses is approximately 396.78 square feet. Some homes don't have a basement area and the maximum size of a basement is 1,134 square feet. The average traffic count of the site is approximately 20,221 cars per day and the maximum is 36,000 cars per day. However, some areas don't have traffic.

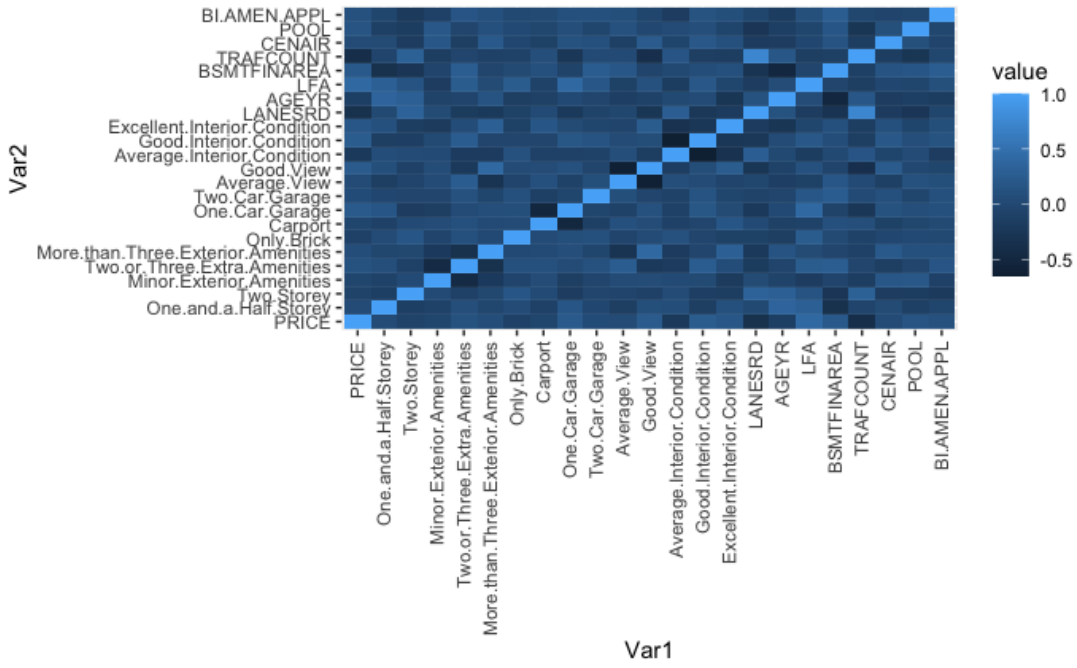
3. Methodology and Model

We must choose the most relevant control variables influencing the dependent variable. Our goal is to determine variables potentially affect the dependent variable, the house price. Therefore, we need to find the most relevant and least relevant variables to avoid the omitted variable bias. Thus, we want to look for pertinent variables that correlate with the treatment variable but are not influenced by it. First, we must include the variables already proven to be related to the outcome and the variables that have a significant effect. We are also looking to exclude variables with many missing values or low variability, variables highly correlated with other variables in the model (causing a collinearity risk), and all the variables that are not linearly related to the outcome.

Moreover, the sample is relatively small. We only have 104 observations to build a Multiple Regression Model. To ensure the model is parsimonious, we limited the variables to put in the model to eight.

First, based on the correlation matrix, we can see that Traffic Count and Lanesrd (Four-lanes Road) correlate negatively with the price. In contrast, LFA (Lot frontage area) correlates significantly with the price. The interior is slightly related to the Price, as in the heatmap below; we can see the Average.Interior.Condition variable is darker compared to other variables.

Figure 1: Correlation Matrix Heatmap



By intuition, a house with a large lot frontage area will have more space, which will positively impact the price. We do not consider house type in our model because in our data it's divided by the number of stories or floors and a house with a large front yard may already have these properties. This also applies to the garage type; we consider the view as an element that affects house prices. Age may be the factor that affects house prices. However, we believe there are more critical factors in the price than age. Instead of age, we consider the interior condition of the model since we believe that interior condition is the age that affects the price. The final variable we count for in the model is basement area; this variable shows the house's size or usable space, impacting the price.

To avoid multicollinearity, we will separate the traffic count and four-lane road into two different models. As in the correlation matrix, we can see a high correlation between the two variables.

From the method above, based on our data, we decided to go with the following variables:

- House View:
  - Average View
  - Good View
- Interior Condition:
  - Average Interior Condition
  - Good Interior Condition
  - Excellent Interior Condition
- LANESRD (Four lane road) (for model 1)
- LFA (Lot frontage Area)
- BSMTFINAREA (Basement Area)
- TRAFCOUNT (Traffic Count) (for model 2)

In this model, we use statistical methods, such as multiple regression, referred in the industry as Automatic Valuation Models (AVMs). This incorporate specific real estate and property attribute data to estimate market value. Unlike the traditional methods approach, which looks at the price of similar properties in the marketplace, based on that property, the price has been decided, which creates high chances of subjective bias.

4. Result

Table 2: Regression Results

The sample has 104 observations of the sales price of residential properties on and around Springbank Drive from January 1998-May 2003.

	Model 1		Model 2	
	Estimated Coefficients	Standard Error	Estimated Coefficients	Standard Error
Intercept	73487.21	12953.91 ***	89405.56	14915.32 ***
House View:				
Average.View	10720.58	6282.63 .	8688.04	6200.86
Good.View	16357.92	7882.73 *	11739.51	7983.41
Interior Condition:				
Average Interior Condition	9103.04	8394.17	8008.40	8199.05
Good Interior Condition	16833.11	8811.48 .	15849.70	8639.34 .
Excellent Interior Condition	20335.26	10768.30 .	19281.56	10555.59 .
LANESRD (Four lane road)	-5942.80	5746.70	-	-
LFA (Lot frontage Area)	39.28	9.34 ***	38.21	9.11 ***
BSMTFINAREA (Basement Area)	13.94	8.71	14.76	8.18 .
TRAFCOUNT (Traffic Count)	-	-	-0.67	0.31*
R-Squared	0.35		0.38	
Adjusted R-Squared	0.30		0.33	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

To determine the appropriate compensation for the house owner, we have created the above two models.

By looking at the negative estimated coefficient value of a four-lane road, we believe if we have a house near a four-lane road, the price of the house will drop an average of 5,942.80 CAD with the average view, good view, and average interior condition, good interior condition, excellent interior condition, frontage area, and basement area constant. ($p < 0.1$)

The estimated coefficient value of the lot frontage area is 39.28 CAD, which indicates that the one square foot increase in the frontage area which increases the sales price on average by 39.28 CAD keeping other variables constant. ($p < 0$)

By looking at the negative estimated coefficient value of traffic count in model 2, we analyze that every additional car on the road will increase the house's sales price on average by 0.67 CAD, keeping other variables constant. ($p < 0.01$)

The estimated coefficient value of the lot frontage area is 38.21 CAD, which indicates that the one square foot increase in the frontage area increases the sales price on average by 38.21 CAD keeping other variables constant in model 2 at ($p < 0$)

From (Table 2) model 1 and model 2, sales price of the house increases on average ($bgoodview = 5637.34 CAD), and model 2 ($bgoodview = 3051.47 CAD) respectively, keeping other variables constant. Similarly, the premium associated with excellent interior condition relative to an average interior condition for model 1 ($bExcellentInteriorcCondition = 11232.22 CAD), model 2 ($bExcellentInteriorcCondition = 11273.16 CAD) by keeping other variables constant.

Overall, by looking at Table 2, we conclude that house prices go up if there is less traffic near the house, a road with fewer lanes, a good view, excellent interior condition, more frontage, and a basement area. The estimated price variation has 30% for model 1 and 33% for model 2, resulting from the variables.

Figure 2: Standardized residual plot of Model 1

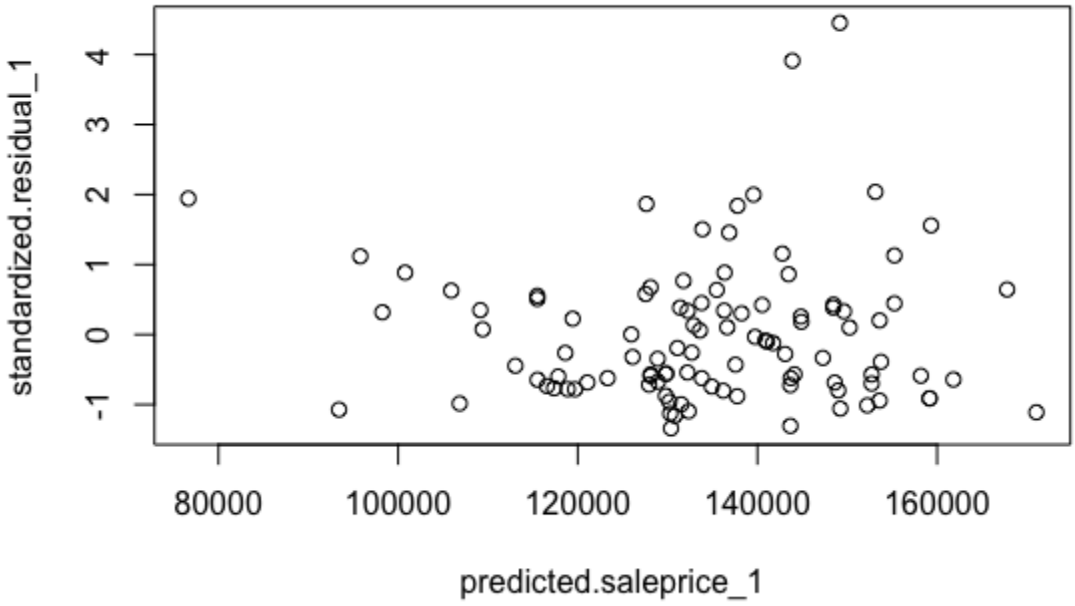
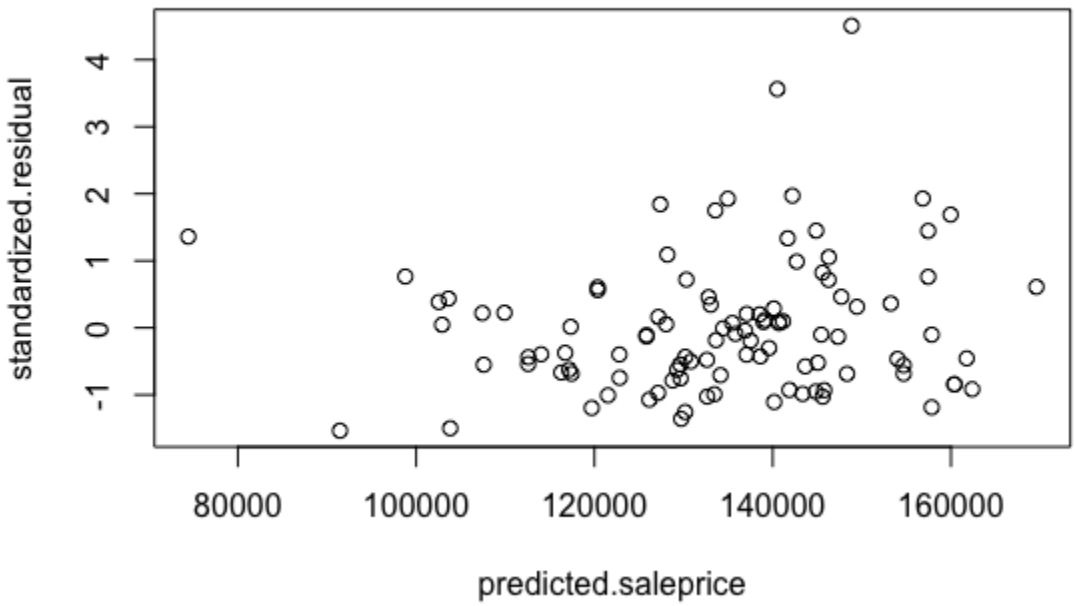


Figure 3: Standardized residual plot of Model 2



5. Implications

1. Expression for model 1:

   LIV = (LFA * B1) + (B2)

   LIV = Loss in value

   LFA = Lost Frontage Area

   B1 = Lot Frontage Area Coefficient of Beta

   B2 (Depreciate value) = Coefficient of Beta of Four lane road if house is on two lanes road otherwise 0

   Total Compensation = LIV + Lump sum amount

2. Expression for model 2:

   LIV = (LFA * B1) + (TFC2 –TFC1) * B2

   LIV = Lost in value

   LFA = Lost Frontage Area

   B1 = Lost Frontage Area Coefficient of Beta

   TFC2 = Increased Traffic Count of cars

   TFC1 = Previous Traffic Count of cars

   B2 = Coefficient of Beta of Traffic Count

   Total Compensation = LIV + Lump sum amount

Table 3: Suggested compensation examples

Property#	Four-lane roads	Current TFC	Increase TFC	Lost Frontage	Model 1 (CAD)	Model 2 (CAD)
1	0	21000	12000	15	18532.00	20613.15
2	0	24000	9000	16	18571.28	18641.36
3	1	26000	7000	13	12510.64	17186.73
4	0	24000	9000	18	18649.84	18717.78
5	0	33000	0	17	18610.56	12649.57

To find the suggested compensation rate for the City of London to the homeowner at Springdrive Bank, we developed two ways to tackle the problems based on the findings above. First, we use the four-lane road to estimate the cost, and second, we use traffic count. For the first model, we suggest that 5,942.80 CAD compensate the homeowner for the house connected with the two-lane road upgraded to a four-lane road plus 39.28 CAD for each foot of lot frontage lost plus the 12,000

CAD lump sum. The alternation suggestion is 0.67 CAD for each traffic count increase to 33,000 plus 38.21 CAD for each foot of lot frontage lost plus the 12,000 CAD lump sum.

6. Limitations

The above analysis has a few limitations in determining the appropriate compensation for the house owner, such as the data set taken for analysis may only include some relevant variables which affect house price. Also, the sample size is relatively small because we have only 104 past sales price of houses sample data from January 1998 - May 2003, and the samples that have been collected may not be random.

7. References

David Picken, B. I. (2015). Chapter 4 The Relationship between Building Height and Construction Costs. In B. S. Herbert Robinson, *Design Economics for the Built Environment.* Chichester, West Sussex, United Kingdom: John Wiley & Sons, Ltd.

Frost, J. (n.d.). *Model Specification: Choosing the Best Regression Model* . Retrieved November 27, 2022, from Statistics By Jim: https://statisticsbyjim.com/regression/model-specification-variable-selection/

Gomez, J. (2022, June 4). *8 critical factors that influence a home's value*. Retrieved November 27, 2022, from Opendoor: https://www.opendoor.com/w/blog/factors-that-influence-home-value

NerdWallet.com. (2018, April 2). *Love a home with a view? How much more you could pay*. Retrieved November 27, 2022, from boston.com: https://www.boston.com/real-estate/spring-house-hunt/2018/04/02/how-much-does-view-increase-home-price/

Nikolopoulou, K. (2022, October 30). *What Is Omitted Variable Bias? | Definition & Examples*. Retrieved November 27, 2022, from Scribbr: https://www.scribbr.com/research-bias/omitted-variable-bias/

Word Count: 1492 (excluding tables and references)