Springbank Drive Final Case Analysis

To City of London Officials,

Thank you for choosing Canning Consultants Inc. We value our clients' trust and confidence in our services and deliver the best results. Hence, before we estimate the loss of market value of each property, we would like to provide the background on the proposed road expansion project and the importance of compensating homeowners appropriately. We have also attached the complete technical report, which contains the methodology and model that we used to estimate the property values to determine the adequate compensation amounts, results, and implications for your reference.

Background and Motivation

London, a Canadian city in southwestern Ontario, is home to the University of Western Ontario, bringing the heaviest traffic of 24,000 cars per day on its Springbank drive. The approved construction project includes widening Springbank drive from two to four lanes between Horton Street and Wonderland Road, and the primary concern is estimating the compensation to the affected properties' owners. Property owners may dispute the compensation amount offered by the government, and the subject property cannot be taken for road expansion use unless adequate compensation is provided in return, which may cause a delay to the project. Therefore, it is essential to provide adequate compensation amounts, considering all factors, including expropriated front area, increased traffic volumes, and other problems like the decline in property values, to ensure the project benefits both government and residents.

It has been a pleasure working with you. Thank you.

Sincerely,

George Canning
Canning Consultants Inc.

Technical Report

Descriptive Statistics

Table 1 below summarizes the descriptive statistics for variables selected in multiple linear regression models for the sample of 104 residential properties drawn from on and around the Springbank drive area from January 1998 to May 2003. The sample data included the details of the past sale prices and 14 other conditional attributes for each property.

Table 1 – Descriptive Statistics

Variables	Mean	Standard Deviation
Price (in CAD)	134470	27744
Traffic Count (in num)	20221	8369
Number of Lanes	0.33	0.47
Frontage Area (in sq.ft)	902.85	255.55
Two Storey House	0.15	0.36
Average Street View	0.58	0.50
Good Street View	0.22	0.42
Good Interior Condition	0.34	0.47
Excellent Interior Condition	0.12	0.32
Basement Finished Area (in sq.ft)	396.78	287.16

From the table1, we observe that about 33% of the houses were located on a four-lane road before the proposed road expansion, while 67% were on a residential street, and the traffic count average mean is 20,221 (no.s), which implies that nearly half of the properties are in heavy traffic area. It is also interesting to note that most properties from the sample are relatively inexpensive homes. For example, based on the interior condition, about 88% of the houses had good/average/fair conditions, while only 12% of the properties had excellent interior conditions; and only 15% of the homes were two-storey while the other 85% are one-and-half/one-storey. However, about 80% of the houses have average or good street views, which makes the property more desirable.

Methodology and Model

To determine the fair value of the property, many characteristics associated with the property play critical roles. For instance, a house with excellent interior conditions will enhance its worth more than a house with average or fair conditions since it becomes more attractive to prospective buyers. Therefore, using the traditional method to value property by comparing the price of similar properties on the market is not a preferred method to accurately capture home values because essential home characteristics should be considered. As a result, Canning Consultants decided to use a quantitative approach (Automatic Valuation Model) to estimate appropriate home values by developing a multiple regression model so that London city residents can be awarded compensation fairly. In contrast, subjective bias is always possible in traditional methods that involve comparing similar properties.

After comprehensively examining all factors related to the house price, we decided to include the following variables in the multiple regression model: Traffic Count, Number of Road Lanes, Frontage Area, Streetscape, House Interior Condition, House Type, and Basement Finished Area.

The variable that has the most significant adverse effect on house prices is traffic. The higher the traffic volumes, the lower the house value since the properties in areas with heavy traffic becomes less desirable. Another variable that negatively affects house prices is the number of road lanes since the increase in lanes will also increase the traffic, decreasing property value. However, as traffic counts and the number of lanes are highly correlated, causing multicollinearity issues, which will reduce the precision of the estimated coefficients, we've considered two models for estimating prices with traffic count and road lanes as independent variables, respectively.

Unlike the variables above that negatively affect house prices, frontage area and streetscape are vital variables that enhance the property value. Generally, a property with a wide frontage area is more desirable and attractive to homebuyers because it can be used for multiple purposes and increases the home's value in the future. In addition, houses with good street views tend to be more expensive than houses without.

Finally, interior condition, house type, and basement finishing area increase house prices. For example, a house with very nice interiors usually positively impacts the house's value because the materials used in the interior are expensive and of better quality. Furthermore, homes with two floors or basements with a finished area are more prominent and spacious than houses with one-and-half floors, providing additional bedrooms and bathrooms, likely increasing the home value. Our model did not include the property's age because of the counterintuitive results it produced when included. Also, the garage is a relevant variable for price estimation, but we excluded it due to concerns about the error term distribution. Please see below our outcome of the multiple regression model using selected variables.

Results

Table 2 – Estimated Property Price Equations by Traffic Count and Lanes Sideroad
Dependent Variable: Estimated Price (in Canadian dollars)

Parameter Estimate	Model 1#			Model 2#		
	Estimate		Std. Error	Estimate		Std. Error
Intercept	96442.27	***	(13675.07)	79478.91	***	(11765.08)
Lot Frontage Area	36.00	***	(9.29)	37.91	***	(9.58)
Traffic Counts	-0.78	*	(0.32)			
Number of Lanes				-6199.53		(5908.09)
House Type						
Two Storey	6692.17		(6953.25)	3897.94		(7066.66)
Streetscape						
Average View	9731.58		(6130.73)	12051.88	.	(6219.87)
Good View	12803.74		(7918.94)	17988.32	*	(7819.02)
Interior Condition						
Good Interior Condition	9797.66	.	(5172.75)	9582.82	.	(5378.57)
Excellent Interior Condition	13091.07	.	(7730.05)	12854.65		(7978.94)
Basement Finished Area	17.82	*	(8.27)	16.50	.	(8.73)
Sample Size	104			104		

Degrees of freedom	95	95
R-Squared	0.3804	0.3484
Adjusted R-Squared	0.3283	0.2935
F-Statistic	7.292 ***	6.348 ***

Note: Bold font (*) indicates that coefficients are statistically significantly different.

*. p < .10, *p < .05, **p < 0.01, ***p < 0.001.*

Model 1 included 'Traffic counts' as one of the independent variables, whereas Model 2 included 'Number of Lanes' as an independent variable for property price estimation, with other variables remaining the same.

As shown in the table. 2, both models have a positive relationship between the frontage area and the property price. This means that by controlling for the traffic counts/number of lanes, house type, street view, interior condition, and finished basement area included in the model, every additional square foot of frontage area is associated with a $36 (Model 1) / $37.91 (Model 2) increase in predicted price. This makes sense because a larger front yard would make a house more desirable. The coefficient on 'Traffic counts' is negative, which indicates that controlling for the other variables in the model, every additional traffic count is associated with a $0.78 reduction in the predicted property price. This intuitively makes sense because most people prefer to live in a neighborhood with less disturbance. The same is true for the variable 'Number of lanes' because additional lanes result in increased traffic. The coefficient on 'Number of lanes' is negative, which indicates that the reduction in predicted price associated with a property on a four-lane road relative to a property on a typical residential location is $6199.53, controlling for the other variables in the model.

Table 3 – Statistical Significance

Parameter Estimate	**Statistical Significance**	
	Model 1	**Model 2**
Lot Frontage Area	Statistically significant at 0.1% level	Statistically significant at 0.1% level
Traffic Counts	Statistically significant at 5% level	-
Number of Lanes	-	Not statistically significant at 5% level
House Type		
Two storey	Not statistically significant at 5% level	Not statistically significant at 5% level
Streetscape		
Average Street view	Not statistically significant at 5% level	Not statistically significant at 5% level
Good Street view	Not statistically significant at 5% level	Statistically significant at 5% level
Interior Condition		
Good Interior Condition	Not statistically significant at 5% level	Not statistically significant at 5% level
Excellent Interior Condition	Not statistically significant at 5% level	Not statistically significant at 5% level
Basement Finished Area	Statistically significant at 5% level	Not statistically significant at 5% level

The Adjusted R-squared for Model 1 is 32.83%, which indicates that the estimated MLR equation explains about 32.83% of the variation in the property price with frontage area, traffic counts, house type, streetscape, interior condition, and finished basement area, as independent variables. The Adjusted R-squared decreased from 32.83% to 29.35% when the 'traffic count' variable was replaced with the 'number of lanes' in Model 2. The implication is that the estimated MLR equation in Model 1 provides a better fit for the observed data.

Implications

Based on the two multiple regression analyses (table 2) performed using the attributes of the 104 properties in the Springbank drive area, we estimated the alternative compensation amounts using the below mathematical expressions and presented them in table 5.

Model 1

Compensation = Lump sum value + Value of the lost frontage area + Loss in property value because of increased traffic

= 12000 + 36.00 * Lost Frontage + 0.78 * Increased Traffic Count

Model 2

Compensation = Lump sum value + Value of the lost frontage area + Loss in property value associated with the increased number of lanes

= 12000 + 37.91 * Lost Frontage + 6199.53 * Increase in number of lanes

Table 4 – Features of the Properties

Property	House Type	Age	Lot Frontage area	Exterior Amen.	Exterior Finish	Garage Type	Street scape	Central Air	Pool	Interior Condition	Basement Finished area	Interior Amenities	Lane Sideroad	Traffic Count
1	0	51	833	0	1	1	3	1	0	5	762	0	0	21000
2	0	49	721	1	2	1	5	1	0	3	550	0	0	24000
3	3	26	1240	0	3	1	3	1	0	4	588	2	1	26000
4	1	34	711	1	2	1	3	1	1	3	512	0	0	24000
5	0	40	798	2	4	1	4	1	1	3	753	1	0	33000

Table 5 – Compensation

Property	Lost Frontage (in sq.ft)	Increase in Traffic count (in num)	Change in lanes road	Compensation (Model 1)	Compensation (Model 2)
1	15	12000	1	$21,900	$18,768
2	16	9000	1	$19,596	$18,806
3	13	7000	0	$17,928	$12,493
4	18	9000	1	$19,668	$18,882
5	17	0	1	$12,612	$18,844

Limitations

Although we have done our best to capture the data to estimate an appropriate compensation for the property owner, there are still significant limitations to consider: the multiple regression model we developed is vulnerable to any changes that could happen due to unforeseen circumstances. To illustrate, most home values generally increase over time. However, the house price in the dataset represents the price at which homes were sold between 1998 and 2003, which means that the data lacks to capture house prices that could potentially increase in the future, resulting in underpaying the compensation amount.

Besides discussed limitation, another limitation can be found in the multiple regression itself. When we developed the multiple regression to estimate the compensation amount, we restricted the variables to eight, which are highly correlated with price because of concerns about the error term distribution. Since we relied on the correlation matrix to select variables that we wanted to include in the multiple regression model, we might have failed to include other variables that might affect the price more significantly than others, causing incorrectly estimating the compensation amount. Therefore, we need to consider the abovementioned limitations to ensure homeowners receive appropriate compensation due to the lane expansion construction.

The word count, excluding the tables and R outputs, is 1480.