PRICED TO SELL: HOW LIST PRICE AFFECTS FINAL SALES PRICE IN THE KINGSTON REAL ESTATE MARKET

by

Katerina Gribbin

An essay submitted to the Department of Economics

in partial fulfillment of the requirements for

the degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

August 2019

Acknowledgments

I would like to express my sincere gratitude to my supervisor Dr. Lloyd-Ellis for his incredible guidance and excellent suggestions. I would also like to extend a special thank you to Bahman Kashi, as without him, this project would not have been possible.

I am also very grateful to Jordan and Rachel at Real Estate Reimagined for giving me access to the data for this project. Finally, I would like to thank the rest of the staff at Limestone Analytics for all of their support.

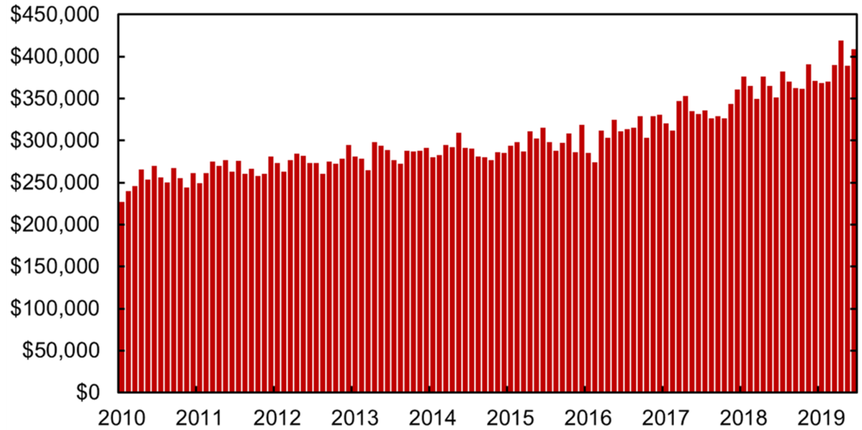
Contents

Introduction

The City of Kingston has an extremely hot housing market, with over 30 per cent of houses in the first half of 2019 selling for above asking price. However, in much of the real estate literature, the asking price is treated as a price ceiling, and an offer at asking price is almost certain to result in an immediate sale. Using a hedonic price regression, this paper finds that underpricing a property by \$10,000 leads to a \$1,600 reduction in the final sales price after accounting for a wide range of property characteristics. Furthermore, conditional upon the potential buyers of an underpriced house engaging in a bidding war, the property sells for \$1,100 less for every \$10,000 by which it was initially underpriced. A dynamic search model is also presented, which predicts that underpricing a property leads to a lower sales price, even conditional upon it experiencing a bidding war. A possible explanation for this result is that the listing price signals how motivated a seller is to sell his property, and that a low price is traded off against a quick sale. Other possible explanations, such as a principal-agent problem between real estate agents and sellers, are also briefly discussed.

Statistics Canada (2018) states that the rental vacancy rate in Kingston at 0.5 per cent — the lowest out of all 35 Canadian cities reported, and roughly half the vacancy rate of Vancouver and Toronto. This hot market can also be seen in the Kingston real estate market. Property prices have risen sharply in recent years (**Chart 1**), and inventory levels are very low (**Chart 3**). Recently, Kingston real estate price growth has outpaced that of Toronto, and far outpaced the rate of inflation (**Chart 4**).

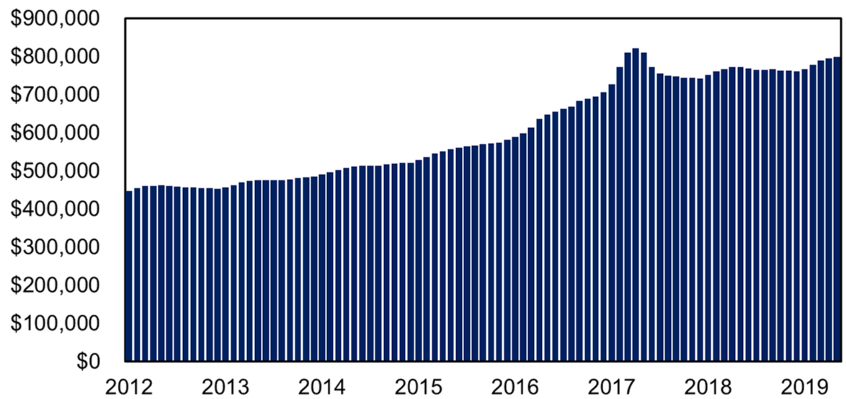
Chart 1: Average Kingston Residential Real Estate Prices



Source: Canadian Real Estate Association

Chart 2: Average Toronto Residential Real Estate Prices

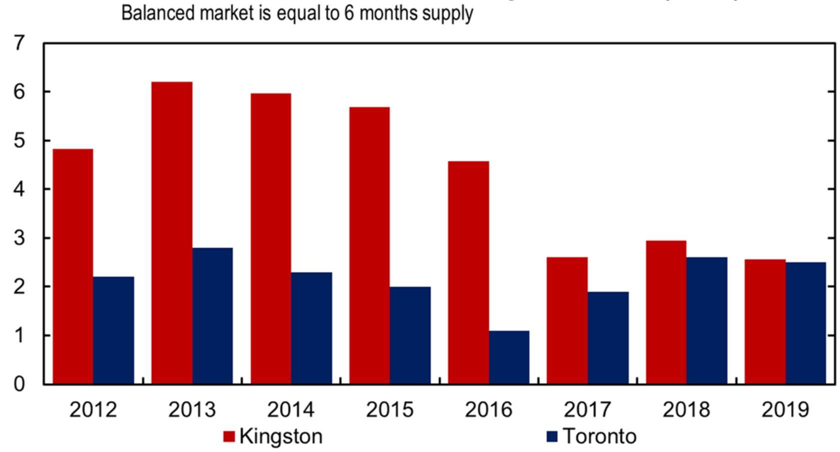
Average composite benchmark price for the city of Toronto



Source: Toronto Real Estate Board

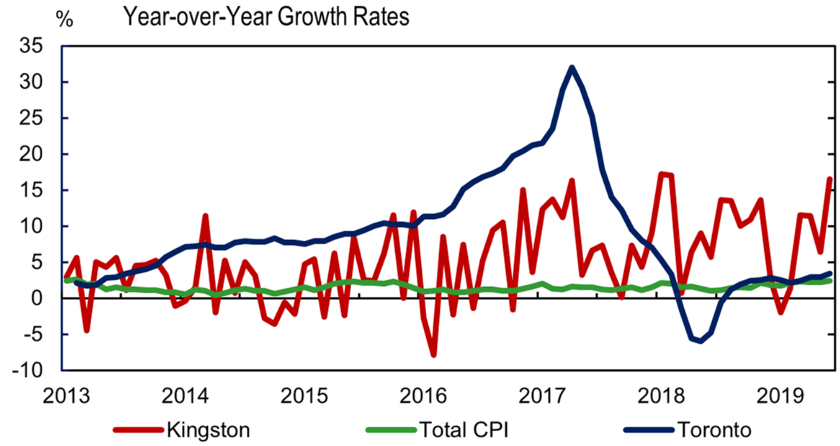
Chart 3: Residential Months of Inventory in Months (June)

Balanced market is equal to 6 months supply



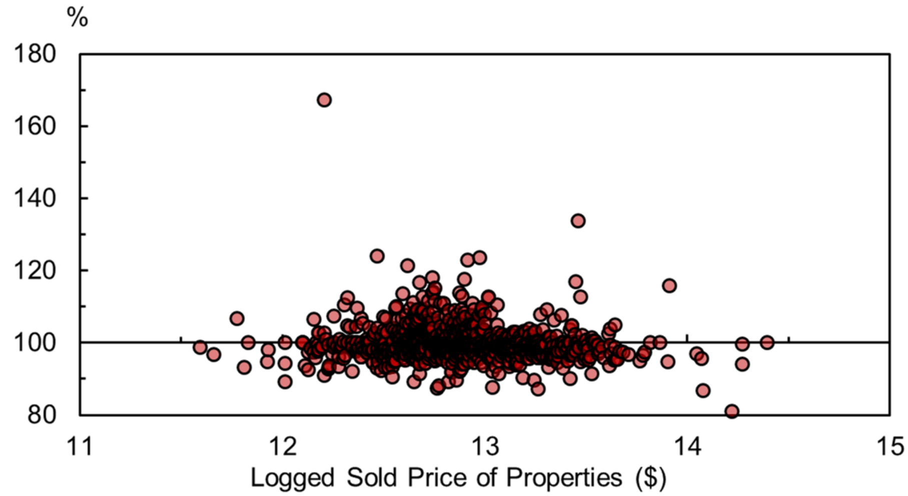
Source: Canadian Real Estate Association and Toronto Real Estate Board

Chart 4: **Real Estate Price Growth in Kingston has Exceeded that of Toronto in Recent Years**



Like many real estate markets in Canada, the Kingston market is decidedly a seller's market. Bidding wars are common, with 31.2 per cent of properties in the dataset selling for above asking price, and a further 12.5 per cent selling exactly at asking price (**Chart 5**). In most real estate markets, sales above asking price are a rare occurrence. While bidding wars are thought to be common in some other Canadian real estate markets — Vancouver and Toronto in particular, Canadian real estate data is surprisingly scarce. No reliable statistics exist on the frequency of bidding wars in other markets. In Toronto, the average house sells at 100 per cent of list price - a number almost identical to that of Kingston. Given this and the recent softness in the Toronto market (**Chart 2**), it is not unreasonable to believe that the percentage of homes that experience bidding wars is similar.

Chart 5: Over 30 Per Cent of Properties Sold for Over Asking Price



Source: MLS

With bidding wars occurring as frequently as they do, it seems likely that some properties are being deliberately underpriced in hopes of triggering a bidding war. While this strategy is commonly recommended by real estate agents in hot markets, the literature is almost entirely silent on whether it is a good strategy. This paper seeks to investigate if underpricing a property leads to it selling for more or for less than its "fundamental valuation," with a particular emphasis on the topic of bidding wars and how they affect the final sales price. As housing makes up approximately three quarters of household wealth in Canada (Statistics Canada, 2018), even modest changes in the selling price of a property can have a significant impact on a given household.

This paper will use a dynamic search model to explore one explanation for why otherwise identical houses can sell for different amounts. In this model, some sellers have a lower reservation price than others. Sellers can use the list price to signal which type of seller they are. Buyers are more likely to visit sellers with low reservation prices and go on to bid less for these types of houses. When multiple bidders bid the asking price, this triggers a bidding war. The results of the model suggest that sellers with a low reservation price sell their homes for less but are more likely to sell. Bidding wars increase the expected sales price for low type sellers but not for high type sellers. Finally, while bidding wars increase the final sales price for low type sellers, this price is still less than what they would have likely received if they had priced their houses higher initially.

In the empirical analysis section, a novel dataset of over 1,000 recent residential property sales in Kingston will be used to examine the relationship between underpricing a property and final sales price. Data has been scraped from internal MLS sold listings, and then matched to neighbourhood characteristics from the Statistics Canada 2016 Census data at the dissemination area level(the smallest unit Statistics Canada reports data for).

In order to find a "fundamental valuation" that is independent of the asking price, a hedonic pricing model approach is used to predict what the property should have sold for based on its observable characteristics. Hedonic modelling is a revealed preference method that considers a good as a combination of characteristics, each of which has an implicit price that buyers have some willingness to pay for. In the housing example, buyers have some willingness to pay for an additional bedroom, a garage and proximity to amenities. Once this fundamental valuation is calculated, it is then possible to determine the extent to which a recently sold property was overpriced or underpriced. I then go on to investigate how list prices are related to this underpricing or overpricing.

Construction of the Dataset

The novel element of this dataset is the inclusion of sold Multiple Listing Service (MLS) listings for the City of Kingston area. As MLS listings for Canadian real estate are generally challenging for academics to obtain, few academic papers have used this data. While a recent court ruling allows real estate agents in Ontario to share MLS data publicly (Competition Bureau, 2018), they are generally hesitant to do so for a number of reasons. Notably, despite the court ruling, the MLS has threatened to sue real estate agents that share sold data publicly (Kalinowski, 2018). Furthermore, real estate agents have little incentive to share data, as access to sold listings is a major reason many buyers and sellers choose to go through agents instead of trying to buy or sell a house themselves. Finally, the software used to access MLS data for this project is extremely old and difficult to work with, making even retrieving the 1,000 listings in this project in Portable Data Format (PDF) form a long and tedious process.

The underlying PDFs contain a wealth of information. An example can be found in Appendix B (where personal identifying information such as address, seller information and real estate agent have been redacted). While the data is extensive, the formatting of the underlying PDFs is very poor. Therefore, zonal optical character recognition (OCR) templates were defined using ABBYY FineReader and applied to each individual PDF. The quality of responses varies significantly between categories and listings. Some important variables, such as square footage, were often left blank. Other variables were not consistently formatted, making data extraction challenging. The underlying dataset contains the complete set of listings sold in the first half of 2019, as well as a small subset of 2018 listings whose sale was finalized in 2019.

MLS listings were also linked to 2016 census demographic data by dissemination area. Kingston contains 196 dissemination areas in total. Most neighbourhoods contain approximately 2000 demographic variables, and there is thus an abundance of neighbourhood variables that can be included.

Hedonic Regression Modelling Literature

What Functional Form Should be Used?

A general hedonic regression is often specified in the following way, where p_n^t is the price of good n in period t, z_{nk}^t is the quantity of characteristic k for good n at time t, and $f()$ is some monotonic transformation of the data (which is generally a natural log, Box-Cox or identity transformation). Often, the independent variable is also transformed, most commonly by taking a natural log. Using the correct specification is important, as incorrect forms will lead to biased and inconsistent estimators.

$$p_n^t = f(z_{n1}^t, z_{n2}^t, \ldots, z_{nk}^t, \varepsilon_n^t)$$

The Kingston MLS dataset is relatively small and encompasses a short window of time. While this is somewhat unusual, the dataset's size and scope is comparable to many previously published papers (**Table 1**).

Chin and Chau (2003) argue the hedonic literature has little to say on the best functional form to choose. However, for datasets comparable to the Kingston MLS data, the literature suggests that simple specifications seem to perform best – particularly in the presence of omitted variable bias. Therefore, this paper attempts a few simple functional forms including linear, log-linear and log-log specifications. Log-type specifications seem to perform best, as they reduce heteroskedasticity in the residuals, allowing a better fit than the linear specification. A short literature review on why this functional form is appropriate follows.

Rosen (1974) wrote one of the early hedonic pricing method papers and was the first to comment on the choice of functional form. He suggests that, a priori, there is no reason to prefer one functional form over another. He goes on to argue for using a goodness of fit criterion (such as the likelihood ratio test) to determine the best form.

Garrod and Willis (1992) state that given the lack of sound theoretical foundations in the hedonic pricing literature, a simple functional form seems to be a reasonable choice. They argue that a semi-log specification has been most widely used in the literature, and that while other methods that employ more complex functional forms can fit the data better, there is not much evidence to support the view that this improves the model. The authors go on to fit a linear, semi-log, log-log and Box-Box transformation to their dataset and find the semi-log form fits best even with a sample size of over 300,000 observations and 40 control variables. Diewert (2003) argues that, at least a priori, a log hedonic pricing model is more plausible than a level specification, as it is generally expected that larger errors will correspond to more expensive items. This is certainly true for the Kingston MLS dataset, and outperforms the linear specification for this reason.

However, simple specifications are not without controversy. One issue is that there is no reason to believe that a hedonic regression should be linear in general. For this reason, some authors have experimented with highly flexible forms such as the Box-Cox transformation. Theoretically, the Box-Cox transformation has some nice properties, as it tends to stabilize variance and transform data to have a distribution that more closely resembles the normal distribution. It can also be used to accommodate a quadratic shape parameter. Rausmussen and Zuehlke (1990) advocate for the use of the Box-Cox specification, as it tends to fit the data very well.

However, the Box-Cox transformation also has a couple of serious downsides. Garrod and Willis (1992) state that "the second-order terms introduced additional multicollinearity problems, which reduced the significance of first-order terms without making any significant improvement to the fit." Cassel and Mendelsohn (1985) make a similar argument and advise against using the transformation.

Linnemen (1980) mentions the transformation cannot be applied to any binary or dummy variables - an important point given that many variables in a hedonic regression are dummies and thus not strictly positive (eg. number of bedrooms, presence of a fireplace, ect.). Note that this is not unique to Box-Cox specifications, as a natural log transformation suffers from the same issue. A Box-Cox specification is also generally very difficult to interpret. This makes arguing for economic significance more difficult.

Another important consideration is that hedonic pricing methods are always missepecified to some extent. In theory, a hedonic pricing model should contain all variables that are costly to produce and generate utility. However, as stated in Butler (1982), given the complexities of the real estate market and data limitations, this is simply not possible. As a result, some amount of omitted variable bias is inevitable. Moreover, as characteristics tend to cluster (eg. houses generally do not contain 7 bedrooms and 1 bathroom), collinearity is also a concern. Butler goes on to state that based on studies that have directly examined which functional form works best, there is no consensus.

To investigate how large a problem misspecification is, Butler estimates two models – one with 4 independent variables and another with 10 independent variables. He finds the misspecification effect to be small. The obvious shortcoming of this exercise is that he does not have the true specification, so both models are misspecified – one is just more misspecified than the other. Moreover, only two of the six variables that he adds to the second regression are statistically significant. As a final point, the adjusted R-squared values that he documents are very low for hedonic pricing regressions for real estate, at only 0.6104 and 0.6498 (where most studies contain an R-squared/adjusted R-squared values of around 0.8 or more), strongly suggesting that either St. Louis has an unusual housing market with a lot of variation, or that his model is missing many important variables.

Cropper et al. (1988) was the first paper to investigate how the choice of functional form affects omitted variable bias. To do this, they use a simulation study where consumers with known utility functions bid for houses with given attributes. When all attributes are included, flexible models, such as the quadratic Box-Cox model, perform best. However, when variables are proxied by other variables, simpler forms (such as linear, log-linear, log-log and linear Box-Cox specifications) perform best. To construct the simulation, the authors randomly select sold MLS single-family home listings in Baltimore City or Baltimore County between 1977 and 1978. Neighbourhood data comes from the 1980 census. Many papers published after Cropper et al. use this paper as a justification for using a simple specification.

While the underlying Cropper et al. study is interesting, it does have several important limitations. Only twelve independent variables are included – hardly enough to convincingly argue that no misspecification occurs in the larger model. Moreover, misspecification is added by replacing square footage with the number of rooms. While proxies are a potential source of misspecification, misspecification may arise for many different reasons, such as failing to include a relevant regressor.

Furthermore, likely due to the age of the Cropper et al. study, the simulations themselves are not up to a modern standard. Only a handful of Monte-Carlo simulations are performed, and each simulation contains a sample size of only 200 observations.

Kuminoff, Parmeter and Pope (2010) revisit Cropper et al.'s influential paper. They argue that the hedonic pricing literature has become increasingly concerned with misspecification since the paper was published. A particular area of concern is that neighbourhood characteristics are often poorly captured, as good data is often not available.

The authors use Monte Carlo simulations and include 2,000 observations per simulation. They find that the addition of spatial fixed effects significantly reduces the omitted variable bias in the cross-sectional data. The effect that they found was larger than the gain from using simpler functional forms. The reduction in omitted variable bias is actually large enough that flexible Box-Cox specifications outperform other simpler forms, such as linear, log and semi-log specifications.

While this seems like an ideal specification, the dataset for Kingston does not lend itself particularly well to fixed effects, as it does not contain much of a time dimension and has no repeated cross sections. Therefore, based on the available literature, a simple functional form seems to be the most appropriate approach, and is the one that is adopted in this paper.

Table 1

Features of empirical hedonic studies: 1988-2008[a]

	Paper Specification	**Literature**
Functional form		
Main specification: lin-lin, log-lin, log-log	•	80%
Main specification: Box-Cox		17%
Sample Size		
Median # of observations	1042	1917
Published in 1989-1998		593
Published in 1999-2008		2459
Distribution of studies by # of observations		
0-200		6%
201-500		10%
501-1000		23%
1001-10,000	•	39%
More than 10,000		22%
Market dimensions		
Geography		
Smaller than a city		7%
City or county	•	42%
Multiple cities or counties		42%
Nation		9%
Time period (#)		
0-1 year	•	7%
1-2 years		22%
2-5 years		26%
5-10 years		28%
More than 10 years		17%

[a]Set of 69 hedonic price studies for property values between November 1988 and November 2008. Reproduced with minor modifications from Kuminoff, Parmeter and Pope (2010)

Theory of Hedonic Pricing

Rosen's (1974) seminal paper on hedonic pricing lays out a description of the theoretical underpinnings of the method. A brief summary of his work will be presented in this section, although it is presented for general interest and is not required to support any other part of this paper.

Rosen argues that hedonic pricing assumes that a good is completely described by its characteristics, and that these characteristics can be perfectly measured. Define z to be the hedonic good and let this good have n observable characteristics. Good z can therefore be written in vector form, where z_i is the quantity of each characteristic that the good possesses.

$$z = (z_1, z_2, \ldots, z_n)$$

The underlying theory assumes that a spatial equilibrium exists, and that this results in a full set of implicit prices for the underlying good. For this reason, hedonic pricing methods are a form of revealed preference, as the implicit or shadow prices guides both consumer and producer decision making.

Each good, z, is valued only for the additional utility it generates for consumers. Rosen argues that the implicit prices the model generates are identical to the hedonic prices, and that these prices are revealed directly by observing the price of different goods and the amounts of each characteristic that are associated with each of these heterogeneous goods. Hedonic prices can be found directly by regressing the price of products on product characteristics. Implicit prices, which are equal to hedonic prices, can be written in the following way:

$$p(z) = p(z_1, z_2, \ldots, z_n)$$

Rosen assumes that no individual buyer nor seller can influence the price of any good. In this way, the market is “thick.” Moreover, the model assumes that $p(z)$ is the market clearing price, and that it is “fundamentally determined by the distribution of consumer tastes and producer costs.” Producers can modify the goods they sell in order to better appeal to buyers, although this transformation is costly. The good, z, is objectively measured in that it has the same attributes for any buyer, although each buyer may value its characteristics differently.

Generally, it is assumed that there are so many goods available that a “spectrum of products” exists, although Rosen admits that this is often not a particularly realistic assumption. Furthermore, $p(z)$ need not be linear. The exception to this is when certain types of arbitrage are possible. However, when an “indivisibility” assumption exists, this type of arbitrage cannot exist. Real estate is generally considered indivisible, as it is not, for example, costless to move a bathroom from one house to another.

Consumers decide on their consumption bundle of z based on their underlying utility function, income, and utility they gain from other goods. Based on their underlying utility function and bid function, an indifference surface is generated. The producer’s decision-making is symmetric, where the goal of maximizing profits replaces the goal of maximizing a utility function, and an offer function replaces the bid function. The optimal choice requires the marginal cost of each characteristic to be equal to the price. Hedonic prices occur where the bid functions of buyers and the offer functions of sellers are tangent to one another. The observations $p(z)$ are the “joint envelope of a family of value functions and another family of offer functions.” However, in of itself, this joint envelope reveals almost no information about the buyers and sellers that generated it. Rosen then goes on to prove that a market equilibrium exists.

Real Estate Literature

One of the key questions this paper examines is why two otherwise similar houses can sell for considerably different prices. Many possible explanations exist, including information asymmetry, different seller types, unobserved heterogeneity in the good, heterogeneity of buyers and sellers, owner equity and principal-agent problems. While this paper does indeed find that two otherwise identical houses can sell for different amounts, the data does not allow for the determination of this reason. Therefore, this literature review provides several possible explanations as to what could be driving the effect and gives potential avenues for future extensions of the paper.

Maury and Tripier (2014) develop a theoretical model to explain how two identical houses can sell for different prices, even in the absence of heterogenous houses. Their result is driven by the fact that home buyers and home sellers can be the same person (i.e. a household sells one house and buys another). In their model, sellers can have one of two strategies: a sequential strategy or simultaneous strategy. Sellers adopting the sequential strategy sell their current house before buying another, while simultaneous sellers attempt to buy and sell a house at the same time and risk not owning a house or having two houses at the same time. As a result, simultaneous sellers can be particularly motivated to sell. Thus, sellers with different levels of motivation to sell their house is a potential explanation.

Glower et al. (1998) offer some empirical evidence for this. Using data from a phone survey asking about seller motivation, the authors find the initial list price conveyed information on how motivated the seller was to sell his property, and that more motivated sellers set lower initial list prices. Sellers having different reservation prices is explored in the dynamic search modelling section of this paper.

However, many other competing explanations also exist. Genesove and Mayer (1997) examine the Boston condominium market in the early 1990s. The market at the time was characterized by declining prices. They find that sellers with higher loan-to-value ratios set higher asking prices, took longer to sell their condos, received higher prices, and were less likely to sell when compared to condo owners with lower loan-to-value ratios. The authors propose the reason for this is that many sellers rely on the proceeds of the sale of their house to be used as a down-payment for their next property. "If one assumes that the unit could be sold at a single 'market' price, the owner would either have to move to a home of much lesser value (because minimum down payments are proportional to housing values) or forgo moving altogether."

Bucchianeri and Minson (2013) conduct a survey of professional real estate agents. When given fictional homes to price, they find that real estate agents suggested underpricing a house in 70.4 per cent of cases. However, "when surveyed anonymously, real estate agents predicted that higher listing prices would lead to higher sales prices, even after we account for individual differences, property fixed effects, and listing time expectations." A possible explanation for this is a principal-agent problem. As real estate agents only receive a small portion of the full sales price, their interests are not necessarily entirely aligned with sellers. They may, for example, prefer to sell quickly at a lower price rather than expend the effort needed to find a buyer willing to pay more.

Levitt and Syvenson (2008) examine the real estate market in the Chicago area. Their hedonic model is in log-log form and uses block-level fixed effects to control for neighbourhood effects. In total, approximately 120 variables are included, with about

60 of these being key word dummies from the property description. They find that when real estate agents sold their own homes, houses took 9.5 days longer to sell, and sold for 3.7 per cent more. They suggest that real estate agents convince their clients to underprice their homes for a quick sale instead of maximizing the expected sales price. Other papers, such as Anglin and Arnott (1991), Geltner et al. (1991), and Rutherford and Yavas (2012) also discuss agency problems in the real estate market. Therefore, this suggests that if sellers are not fully informed, they may unknowingly underprice their homes, and ultimately receive less for their property.

Rutherford, Springer and Yavas (2005) find further evidence of agency problems. They find that when agents sell their own homes, they sell equally quickly and at a 4.5 per cent premium when compared to other sellers. They state that "the general conclusion of the earlier models is that although the percentage commission system ensures the interests of the agent to be in the same direction as those of the client, it fails to align the magnitude of the interests of the agent with those of the client."

Another possible explanation is that buyer and seller heterogeneity may explain some portion of price differentials not accounted for in housing characteristics (e.g. tastes, preferences, patience, search costs, asymmetric information, etc.). While most papers do not have the data necessary to measure these effects, a few papers have examined this.

Cotteleer et al. (2008) account for buyer and seller heterogeneity when looking at agricultural land prices in the Netherlands. Because farmers have a strong preference for buying property near their existing agricultural land, and because markets are thin, a substantial amount of market power can be at play. This can result in excess surplus for either the buyer or seller. They find market power is important, and that as the number of sellers increases, the total selling price per hectare falls. A few other variables, such as the age of the buyers and sellers, as well as the income of sellers are found to be statistically significant as well, which the authors use as proxies for market power.

Lacobini and Lisi (2012) attempt to control for this heterogeneity in a way that does not require information on buyer and seller characteristics. They state that the standard assumption that markets are "thick" is often not true in real estate markets. Thus, buyers and sellers can potentially have market power. Moreover, as gathering information on a property is costly and not always possible, buyers and sellers may have incomplete information. The paper presents a model that is compatible with asymmetric information. The authors then go on to examine empirically if market power exists in the Italian real estate market.

While the authors do find evidence of market power, there are serious statistical issues with their underlying analysis. Essentially, the authors create a dummy for buyer or seller market power whenever the hedonic model poorly predicts the sold price. The authors find that the inclusion of these dummies significantly improved model fit. While this is true, it fails to prove much, as the dummies were created only for observations that were poorly explained by the standard hedonic pricing model. Thus, it is not clear if the dummies are capturing market power or some other omitted variable.

In conclusion, there are many competing theories in the literature as to what drives heterogeneity in real estate prices. In the Kingston market, it is not possible to say which effects are most important.

Commonly Included Hedonic Pricing Variables

In many papers, hedonic regressions using real estate data are relatively sparse and contain only a handful of explanatory variables. This is likely due data limitations for many real estate data sets. While variables such as square footage, number of bedrooms and the presence of a garage are commonly included as they are known to be important explanatory variables, there is much less consensus on other variables – particularly neighbourhood variables.

Sirmans et al. (2005) provide a meta-analysis on which variables are important in a hedonic regression. They state that nine variables are generally agreed to be important: square footage, lot size, age, bedrooms, bathrooms, garage, swimming pool, fireplace and air conditioning. The following table was presented in their paper.

Table 2

Variable	Number of Studies	Mean Effect	Standard Deviation	Minimum	Maximum
Square footage	64	0.33757	0.2313	-0.087	0.92
Lot size	41	0.03334	0.05929	-0.006	0.213
Age	82	-0.00864	0.00773	-0.045	0.011
Bedrooms	45	0.03772	0.08162	-0.082	0.31
Bathrooms	58	0.0875924	0.07265	-0.03	0.32
Garage	31	0.10819	0.07202	0.01	0.243
Swimming pool	37	0.0771445	0.03382	0.011	0.182
Fireplace	35	0.08934	0.05345	0.002	0.232
Air Conditioning	31	0.08347	0.07388	-0.072	0.31

While including these variables is a good starting place, the inclusion of these regressors is not enough to avoid omitted variable bias entirely. The lack of literature on other important explanatory variables makes the choice of final hedonic regression specification more challenging.

Literature on Sold Price, Asking Price and Days on the Market

The central question this paper attempts to answer is how the list price affects final sales price, and if underpricing a house in hopes of triggering a bidding war is a good strategy in order to maximize expected sales price. A small literature exists on the relationship between sales price and list price and finds some evidence that a higher list price marginally increases the final sales price. However, a higher list price is also thought to increase time on the market. Finally, bidding wars are a topic almost entirely ignored by the literature. Moreover, the vast majority of papers do not use data from markets that are as hot as the current Kingston real estate market and may therefore not be a good source of comparison.

Bucchianeri and Minson (2013) attempt to determine how the initial asking price affects the final sales price. They cite two possible competing mechanisms. Firstly, they state that much of the auction literature argues that goods which start at initial lower prices sell for more (i.e. herding behaviour. On the other hand, some literature on price anchoring suggests that buyers anchor their reference point to the first price presented, even if the initial price is inflated, and use this as a benchmark on which to base their purchasing decisions. Using the anchoring mechanism, an otherwise identical but higher priced house would sell for more.

The authors first estimate the value of over 330,000 homes using a hedonic regression from 1988 to 2009. Of this sample, around 14,000 homes were sold multiple times. They then specifically examine houses that were resold. They find that over-pricing a house by 10 to 20 per cent increases the final sales price by $117 to $163. While the effect is statistically significant, it is arguably not very economically significant.

Arguably, the regression run in this paper is most similar to the Kingston hedonic pricing regression. While the direction of the effect is the same in both papers, the Bucchianeri and Minson paper finds a much smaller effect.

Beracha and Seiler (2013) use a very simple hedonic pricing model to find the true value of a property, and examined how "just below" pricing strategies (eg. $249,999), "round" pricing strategies (eg. $250,000), and "exact" pricing strategies (eg. $250,088) affect the final sales price. They find that on average, sellers using a "just below" pricing strategy tend to overprice their home more, and that this strategy tends to allow sellers to sell their homes for the highest price. One element that the authors did not consider is whether the higher sales price is simply a result of sellers overpricing a home more, and not due to the choice of pricing strategies.

Miller and Sklarz (1987) examine the relationship between list price and sales price of condos in Hawaii. They argue that consumers often mistakenly perceive higher priced goods as being higher in quality, even when price and quality are often weakly correlated. They therefore speculate that buyers may perceive a higher priced home as being of higher quality, and this could influence their willingness to pay for the property. They argue that a higher list price increases sales price and time on the market, but do not directly quantify the effects.

Cubbin (1974) attempts to estimate the quality-adjusted price of homes sold in Coventry, England between 1968 and 1970. His focus is on how quality-adjusted price was related to time to sell, and he comes to the conclusion that the higher the price, the faster the home sells. This is an unusual result in the literature. However, he justifies it by arguing that buyers were judging quality based on asking price.

Therefore, based on the small sample of literature available, it appears that increasing listing price should weakly increase the final sales price. This effect is likely being driven by price anchoring effects or because buyers are using price as a proxy for quality.

Merlo and Ortalo-Magné (2004) provide perhaps the best explanation of the relationship between list price, time on the market, and sales price. They use a dataset from England that allows the authors to directly observe a property's final sales price, days on the market, all offers made on the property, and the number of visits by potential buyers. The authors find that the initial list price influences the arrival of offers, which ultimately determines time on the market. As days on the market increase, visits by potential buyers decrease. The longer the property is on the market, the lower the offers are relative to the sales price, the higher the probability that an offer is accepted, and the lower the sales price is relative to the list price. The authors then go on to state that a higher list price may lead to a higher sales price but results in the property spending longer on the market.

The authors specify their hedonic regression in levels and use the initial listing price as their dependent variable. The paper goes on to define the variable ILISTPRES, which is defined as the list price minus expected list price based on the hedonic regression results. Thus, "ILISTPRES captures the extent to which a property is either over-priced or under-priced relative to other, similar properties." About a quarter of houses change the listing price over the sample period (generally, list price is decreased as a result of the property being initially overpriced and spending a long time on the market). The authors go on to use this variable to examine the probability of price adjustments, and find that in hotter real estate markets, turnover is greater, matches occur more quickly and frequently, and buyer's success rate is lower. A similar variable using prices and "fundamental values" of a properties in Kingston will be generated in the empirical section of this paper.

Sirmans, Turnbull and Dombrow (1994) examine if houses that sell very quickly (within 48 hours) are underpriced. Their dataset is from Baton Rouge, Louisiana. Based on their simple hedonic model, they find that this does not appear to be the case. The authors suggest that institutions exist which prevent sellers from systematically underpricing their homes.

Tucker, Zhang and Zhu (2013) examine how the number of days on the market affects sales price. One issue is that in many housing markets, sellers can relist their house if it has been on the market for a long time, making it look like a new listing to potential buyers. This is important, as more days on the market is generally seen as a negative quality signal to potential buyers. The authors examine a policy change in Massachusetts, where a law was passed that prevented this relisting from occurring and compared it to Rhode Island. They find that for houses that were on the market when the policy change occurred, two strategies existed to overcome the policy change: lower the price and hope the house sells faster, or increase the price and hope buyers attribute the slow sale to the high price instead of low quality. They find the policy change resulted in a large effect, and that buyers had generally been poorly informed on the use of this tactic before the policy came into effect. In Massachusetts, 35 per cent of homes were affected by the policy change. The empirical results suggest the policy change ended the relisting tendency, reduced days on the market and decreased house sales prices (where the effect was a \$16,000 sales price cut by houses affected by the policy change).

Asherfelter and Genesove (1992) examine the final sales price of otherwise identical condos that sold by individuals visiting a property, and compare it to condos that sold at auction. Their sample is based on 83 condos in Princeton, New Jersey. They found that an otherwise identical condo sold at a 13 per cent premium at auction. While this is not identical to a bidding war, it may suggest buyers in an auction act irrationally and pay more than they should for an otherwise identical property. Therefore, this might provide weak evidence to suggest that underpricing a house in hopes of starting a bidding war may be a sound strategy.

Theoretical Modelling Exercise

In order to theoretically model how undervaluing or overvaluing a property affects final sales price, some form of search model that matches buyers to sellers is necessary. More importantly, if buyers are modifying their search behaviours as a result of the posted price, this search model must be flexible enough to account for this. This motivates the choice of using a directed search model.

A large literature on dynamic search modelling exists, with real estate being a relatively common application. The model chosen for this exercise needs to have a few underlying properties in order to have any hope of capturing the dynamics of the Kingston market that we are interested in modelling.

First of all, sellers of otherwise identical properties need to post different asking prices. Next, the price of the property needs to directly impact the search made by buyers. Finally, the model needs to allow for the list price to not entail full commitment, and for a bidding war to occur if multiple buyers arrive at a house and are willing to pay over the asking price.

The directed search model by Albrecht, Gautier and Vroman (2009) does a good job at modelling these effects. In their model, two types of sellers exist: sellers with a low reservation price and sellers with a high reservation price. Ideally, sellers would like to signal that they are a low type seller during the initial search phase (as more buyers will visit their property if this is the case), and signal they are the high type seller during the bidding process (as buyers are willing to bid more if they believe sellers are the high type). However, the only signal sellers have is through the asking price. The authors go on to prove that a separating equilibrium exists for this model. In summary, buyers will be able to directly observe if sellers are the low or high type by observing the asking price, and that no seller has an incentive to misrepresent their type.

When searching for a house, buyers randomly choose to visit low or high type sellers with some probability, where all houses for sale are homogeneous. The expected payoff of visiting either seller type must be identical and the probability of visiting each seller type adjusts to guarantee this. Low type sellers post prices between $[0, s)$ and have a reservation price of 0, while high type sellers post prices between $[s, 1]$ and have a reservation price of s. Sellers are indifferent between posting any price in this interval as a direct consequence of payoff equivalence (i.e. any prices in this interval will generate the same expected payoff for the seller). Note that there is no overlap between these two types (as they have no incentive to misrepresent their type), and thus buyers can deduce the reservation price of the seller that they are visiting. Sellers posted price involves only limited commitment in that it is not binding but does signal something. Buyers randomly choose which seller to visit at random with equal probability after deciding which type to visit.

Buyers only visit one property in this model. Once they visit the property, they draw some random variable, x, that determines how good a match the property is. This value will directly influence any bid they make. The random variable, x, is uniformly distributed between zero and one. Buyers bid function takes the following form, where θ is the buyer to seller ratio (where θ generally differs for buyers visiting low type and high type sellers):

$$b(x) = \begin{cases} 0 \text{ for } 0 \leq x \leq s \\ x - \frac{1-e^{-\theta(x-s)}}{\theta} \text{ for } s \leq x \leq x^* \\ a \text{ for } x^* \leq x \end{cases}$$

Therefore, if buyers draw an x that is lower than the seller's reservation price, they will not bid. If buyers draw a value of x that is higher than the seller's reservation price but lower than the value of x that causes them to bid asking price, their bid function takes the form $x - \frac{1-e^{-\theta(x-s)}}{\theta}$. Otherwise, the buyer bids the asking price. In the case that a seller receives multiple offers at the asking price, an escalator bidding war begins, and is equivalent to a second price auction.

Buyers receive the following payoffs for visiting low type and high type sellers:

$$v_L(r;q,\theta) = \frac{1 - e^{-\theta_L} - \theta_L e^{-\theta_L}}{\theta_L^2}$$

$$v_H(r;q,s,\theta) = \frac{1 - e^{-\theta_H(1-s)} - \theta_H(1-s)e^{-\theta_H(1-s)}}{\theta_H^2}$$

The probability buyers visit high type sellers will adjust to ensure that $v_L(r;q,\theta) = v_H(r;q,s,\theta)$. Thus, fewer buyers will visit high type sellers when s increases. Note that θ_L is equal to the ratio of buyers that visit low type sellers divided by the total number of low type sellers, and θ_H is equal to the ratio of buyers that visit high type sellers divided by the total number of high type sellers.

The paper does not give any guidance on the appropriate parameter values to choose. As a result, parameters have been approximately calibrated from Kingston MLS data.

First of all, a ratio of low type to high type sellers is must be chosen. While this ratio is unknown, there are a few methods that can be used to estimate its value. As small changes to this variable have a relatively minor effect on the simulation results, a rough approximation should be sufficient. To do this, I calculate what percent of homes sell for less than their list price to approximate the number of homes that are overvalued. This corresponds to a value of 56.3 per cent. I take this to be the number of high-type sellers.

Another parameter that is needed for the model is the ratio of buyers to sellers. This information is unavailable for the Kingston market, and it is difficult to come up with an educated guess as to what this ratio is. Given the lack of inventory and frequency of bidding wars, it seems reasonable to assume that there are more buyers than sellers. A value of 2.5 was ultimately chosen, as this value tends to result in bidding wars occurring roughly 30 per cent of the time in the model – similar to the fraction of bidding wars in Kingston.

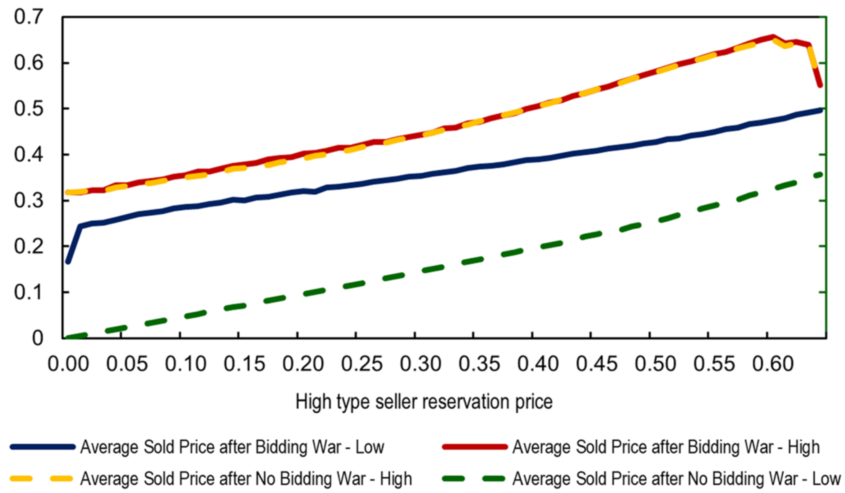
Finally, a value for the high type seller reservation price must be chosen, where low type seller reservation prices are normalized to zero. This information is unavailable, and thus it was simulated over a wide range of values. Values above 0.64 are not simulated, as they result in buyers never visiting high type sellers.

The model suggests that low type sellers receive significantly less for their homes (**Chart 6**). While low type sellers prefer to experience a bidding war, as it increases the expected sold price, it still sells for less than the houses of high-type sellers. Interestingly, while bidding wars increase the expected sales price for low type sellers, they do not increase the expected sales price for high type sellers. Furthermore, home prices are strictly increasing in the high type seller's reservation price up until a reservation price of 0.60.

Parameter values calibrated for the Albrecht, Gautier and Vroman (2009) dynamic search model	
Number of buyers	2500
Number of sellers	999
Number of low type sellers	437
Number of high type sellers	562
Low type seller reservation price	0
High type seller reservation price	Varies

Chart 6: Average price of low and high type homes that experience bidding wars

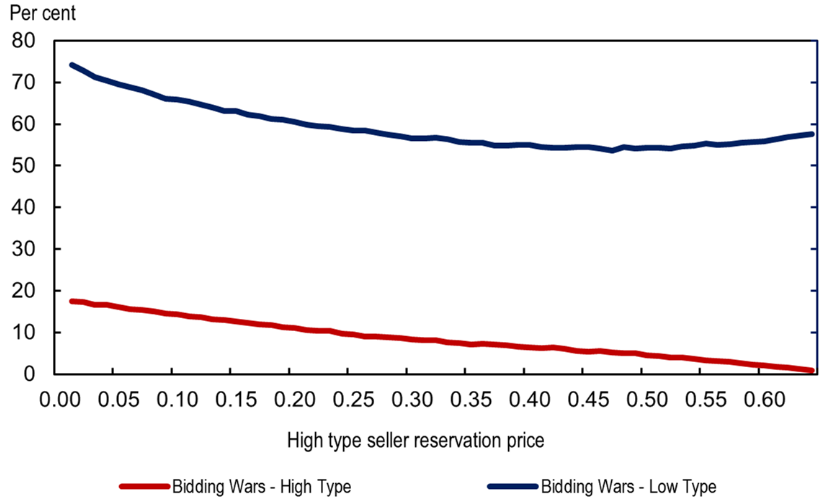
Simulation results of 2,500 buyers, 436 low-type sellers and 563 high-type sellers



Low type sellers are significantly more likely to experience bidding wars on their properties than high type sellers (**Chart 7**). High type sellers become increasingly less likely to experience bidding wars as their reservation price increases. On the other hand, low type sellers' rate of bidding wars is U shaped. The reason for this is that there are two competing effects. As s increases, more buyers choose to visit low-type sellers, and this increases the probability of a bidding war. However, recall that low type sellers post prices between $[0, s)$. Thus, a larger s increases the range of prices low-type sellers can post while still signaling that they are the low type. As a low type seller is indifferent between any list price between $[0, s)$, as they receive the same expected revenue either way, a larger s makes it more likely that a low type seller will post a relatively high price, and thus reducing the probability that two or more buyers will bid the list price.

Chart 7: Per cent of homes sold that experience bidding wars

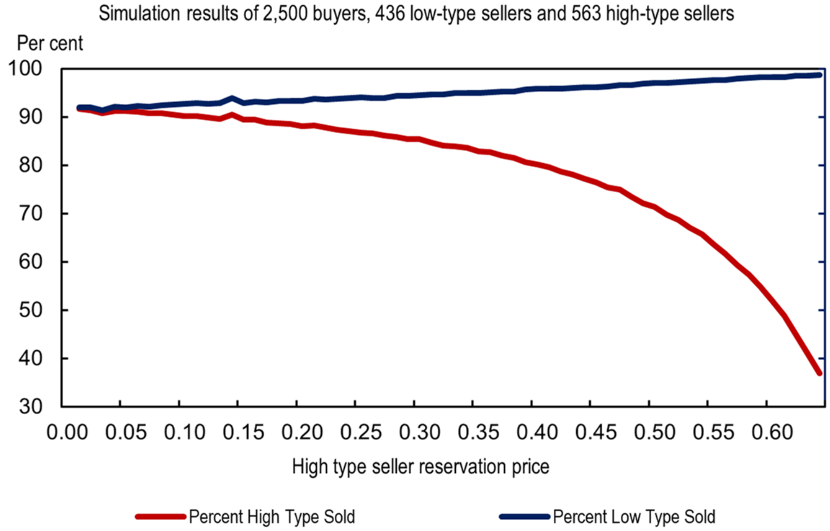
Simulation results of 2,500 buyers, 436 low-type sellers and 563 high-type sellers



Low type sellers are more likely to sell their homes than high type sellers (**Chart 8**). As s increases, low type sellers become increasingly more likely to sell their homes, while high type sellers become increasingly less likely to find a buyer willing to pay their reservation price. This is driven by the fact that buyers become more likely to visit low type sellers as s increases. The overall volume of sales decreases as s increases, as there are fewer buyers willing to pay at least the reservation price for high type homes.

Chart 8: Per cent of homes listed that ultimately sell

Simulation results of 2,500 buyers, 436 low-type sellers and 563 high-type sellers



In conclusion, the model predicts that if the list price conveys information on the seller's motivation to sell, lower priced properties are more likely to sell, are more likely to experience a bidding war, and sells for less on average when compared to higher priced homes.

Characterization of Kingston Real Estate Data

A major advantage of using MLS data to conduct the analysis on the relationship between sold and list price is that it contains almost every property that was sold in the Kingston market over the period under consideration. Only houses that were sold privately or were for sale by owner (FSBO) are not included in the sample.

This dataset contains only residential properties within the boundaries of Kingston. Therefore, vacation cottages and properties within easy commuting distance of Kingston but not within the city limits are not included in the analysis. It should also be noted that the period studied is short – a span of approximately 6 months.

The following table (**Table 3**) contains summary statistics of all variables included in the hedonic regressions that follow. When variables were missing, the variable was dropped from the underlying regression. While this is not without issue, other common methods to deal with missing values, such as imputing the average value of the variable, are also arbitrary in nature and can potentially distort regression results. Of particular concern is the fact that many dropped observations were of newly built properties that cannot be assigned a dissemination area, as the postal code did not exist at the time of the 2016 census, or the property does not yet have property tax data. This has the potential to skew results, but fortunately effects only a small number of observations.

The following chart (**Chart 9**) shows the distribution of property sold prices over the period. While most properties sales price falls between $250,000 and $450,000, the distribution is skewed right (in reality, this skewness would appear even more pronounced if the final bin had not been set to include all properties sold for over $1,000,000).

Chart 9: **Distribution of the final sale price of Kingston real estate**

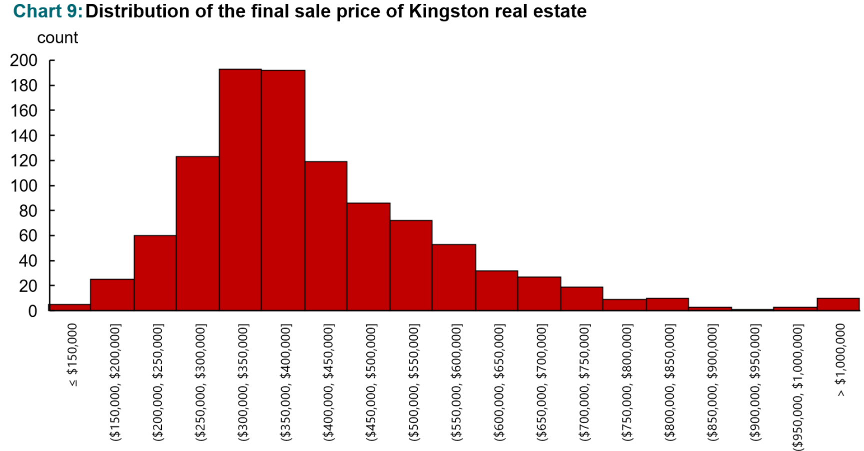


Chart 10: **Sale price as a percent of list price**

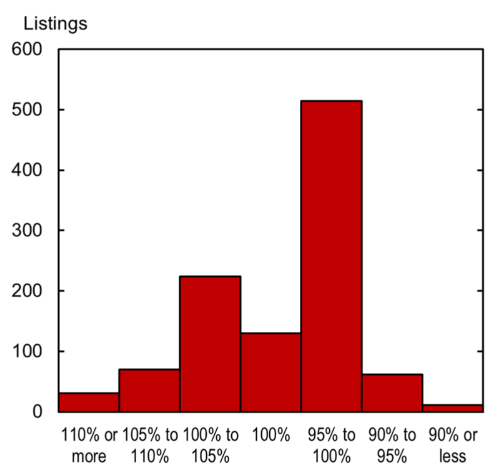


Chart 11: **Properties tend to sell relatively quickly**

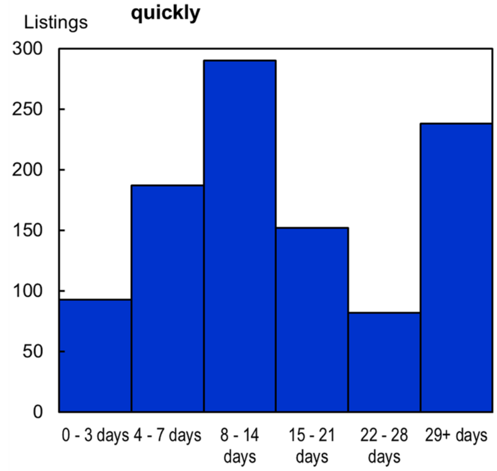


Chart 12: **Property tax assesments undervalue houses, with more expensive homes being systematically undervalued by a larger extent**

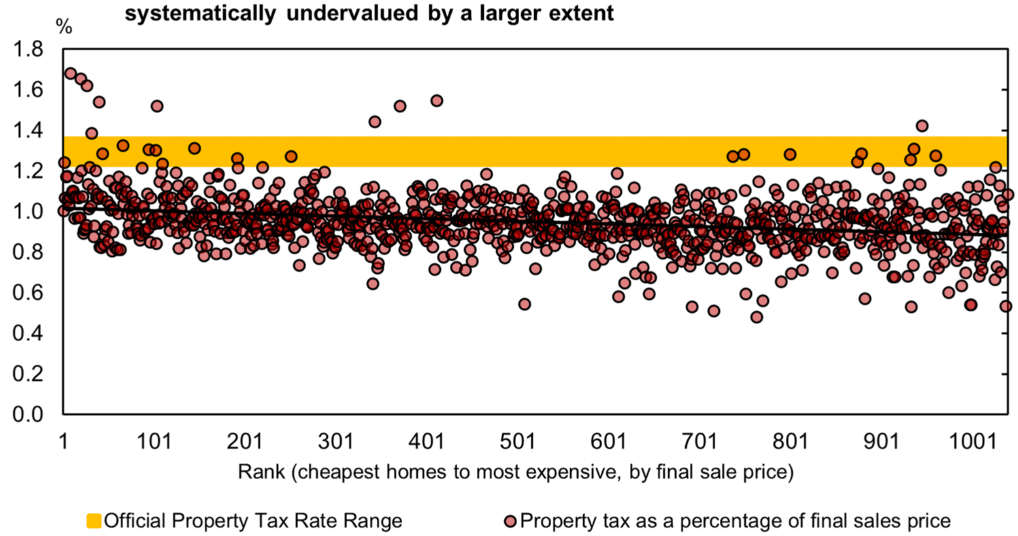


Table 3 - Sample Statistics

	Mean	Median	Standard Deviation	Minimum	Maximum	Sample Size
House Specific Variables						
List Price	$420,782.13	$379,900.00	$178,010.41	$110,000.00	$1,850,000.00	1042
Sales Price	$418,342.65	$382,450.00	$169,842.16	$108,500.00	$1,780,000.00	1042
Sales Price as a Per Cent of List Price	99.89	99.34	4.87	81.08	167.22	1042
Days on the Market	23.72	13.00	33.45	0.00	489.00	1042
Distance From City Hall (km)	7.57	7.46	4.99	0.13	29.92	1042
Property Tax as a Fraction of Sales Price (%)	0.95	0.95	0.14	0.48	1.68	989
Condo Fees	$405.87	$297.30	$272.40	$60.00	$1,192.00	95
Lot Area (sq/ft)	10552.89	4159.48	86086.54	0.00	2280696.00	1042
Square Footage	1722.61	1600.00	575.33	521.00	4381.00	356
Square Footage Proxy	810.90	782.75	222.90	261.12	2758.66	1024
Above Ground Bedrooms	2.98	3.00	0.80	0.00	7.00	1042
Basement Bedrooms	0.39	0.00	0.67	0.00	5.00	1042
Full Bathrooms	1.90	2.00	0.72	0.00	5.00	1042
Half Bathrooms	0.59	1.00	0.56	0.00	3.00	1042
House (dummy)	0.71	1.00	0.46	0.00	1.00	1042
Apartment (dummy)	0.05	0.00	0.22	0.00	1.00	1042
Garage (dummy)	0.74	1.00	0.44	0.00	1.00	1042
Waterfront (dummy)	0.02	0.00	0.14	0.00	1.00	1042
Basement (dummy)	0.90	1.00	0.30	0.00	1.00	1042
Tenant Occupied (dummy)	0.10	0.00	0.30	0.00	1.00	942
Owner and Tenant Occupied (dummy)	0.02	0.00	0.15	0.00	1.00	942
Vacant Property (dummy)	0.17	0.00	0.38	0.00	1.00	942
Fireplace (dummy)	0.42	0.00	0.49	0.00	1.00	1042
New Build (dummy)	0.05	0.00	0.23	0.00	1.00	56
Unknown Age (dummy)	0.16	0.00	0.36	0.00	1.00	162
0-5 Years Old (dummy)	0.07	0.00	0.25	0.00	1.00	71
5-10 Years Old (dummy)	0.09	0.00	0.29	0.00	1.00	56
11-20 Years Old (dummy)	0.18	0.00	0.39	0.00	1.00	191
21-30 Years Old (dummy)	0.12	0.00	0.32	0.00	1.00	120
31-40 Years Old (dummy)	0.13	0.00	0.33	0.00	1.00	132
41-50 Years Old (dummy)	0.07	0.00	0.26	0.00	1.00	78
51-99 Years Old (dummy)	0.10	0.00	0.30	0.00	1.00	104
100+ Years Old (dummy)	0.03	0.00	0.18	0.00	1.00	33
Neighbourhood Specific Variables						
Homes Vacat (%)	5.26	1.41	10.28	0.00	78.45	1007
Median Household Income	$86,060.21	$90,624.00	$27,453.31	$24,512.00	$155,904.00	1007
Median Age	41.84	41.00	8.03	22.20	72.50	1007
Married (%)	48.13	50.00	12.99	11.29	70.71	1007
Less Than High School (%)	12.53	11.13	5.56	0	40	1007
Bachelor's Degree or Higher (%)	29.19	27.05	12.87	0	64.77	1007
Employed (%)	61.59	61.79	9.93	20.22	79.49	1007
Moved in the Last Year (%)	15.68	15.25	7.67	0	62.5	1007

Chart 13: Breakdown of Real Estate Type

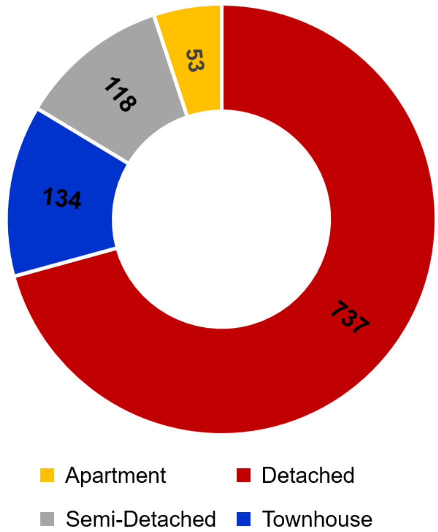
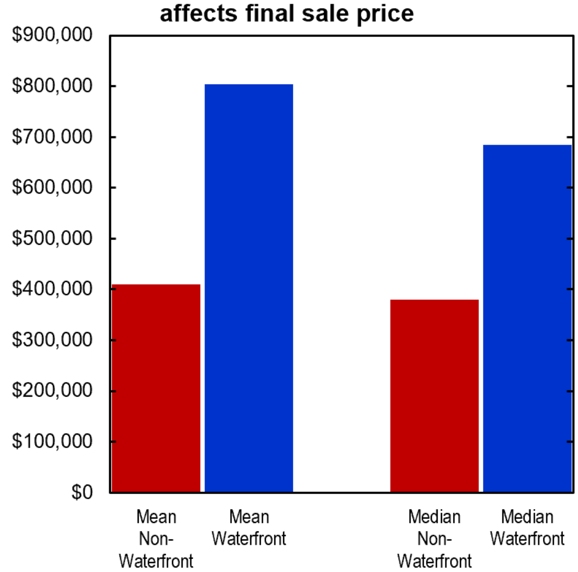


Chart 14: Waterfront property significantly affects final sale price



Data also suggests that properties generally sell close to asking price and relatively quickly (**Chart 10** and **Chart 11**). This is unsurprising given the tightness of the market at this time. The City of Kingston uses property assessment data to determine how much property tax is due on a property every year. In 2018, this rate was pegged at between 1.233 per cent and 1.359 per cent of the assessed value. The assessed values are, in general, much lower than actual sales price for most Kingston properties (**Chart 12**), with higher priced homes paying disproportionately less tax. This suggests that assessed value does a poor job at capturing the actual value of houses in Kingston. Most properties sold in the sample were detached homes (**Chart 13**), with few apartments sold over the period. Finally, waterfront is the variable that has the single largest effect on property prices in the sample (**Chart 14**).

Hedonic Regression Results

This section presents the empirical specification and quantifies the effect that list price has on final sales price after adjusting for a wide range of control variables.

A few variables commonly included in a hedonic regression were problematic, and at times needed to be proxied.

Square footage is a very commonly included variable in hedonic regressions and tends to have a large and positive effect on property prices. Unfortunately, only about a third of properties in the sample have square footage recorded in their MLS listing, and these observations generally corresponded to more expensive houses. Not including square footage would be a major omission and would inevitably lead to omitted variable bias. However, a sample size of roughly 350 properties that are not randomly selected from the full sample of properties is even more problematic. In an effort to reduce the omitted variable bias stemming from omitting square footage, a proxy was calculated. In the vast majority of listings, room measurements are included for several rooms. This is an imperfect proxy, as not all rooms are listed, and some real estate agents are more fastidious about reporting this variable for many rooms than others. While evidently not without issue, this variable is extremely statistically significant, has a large positive effect on property prices in all regression specifications, and is available for almost all properties in the sample. Thus, this proxy has been included in the regression.

The age of property variable is another commonly included variable in hedonic regressions. However, due to the format of the underlying data, age is only available in broad categories (eg. 21-30 years old), and a significant minority of houses have age recorded as unknown. In order to include age in the hedonic regression, 9 categorical variables would need to be included (as 10 categories exist). This specification was tested, and while dummies were significant, a large reduction in the degrees of freedom was required in order to only marginally improve the fit. Moreover, it was not feasible to use the median value of the range as the age, as many properties were listed as unknown in age, and some ranges were over 50 years wide. Therefore, age variables were ultimately excluded from the final specification of the regression .

Three simple hedonic model specifications were run: linear, log-linear and log-log forms. All models fit relatively well, but the log-linear and log-log specifications fit the data better than the linear specification. The R-squared values for the specifications were 0.7726, 0.8326 and 0.8366 respectively – respectable values by the standards of the literature. Charts 15, 16 and 17 show the residuals for the three specifications. The linear specification has a distinctive fan shape, suggesting heteroskedasticity in the residuals. Therefore, it is not surprising that the log specifications in chart 11 and 12 result in a better fit, as log transformations tend to be variance stabilizing.

Chart 15: **Linear Specification Hedonic Pricing Regression Residuals**

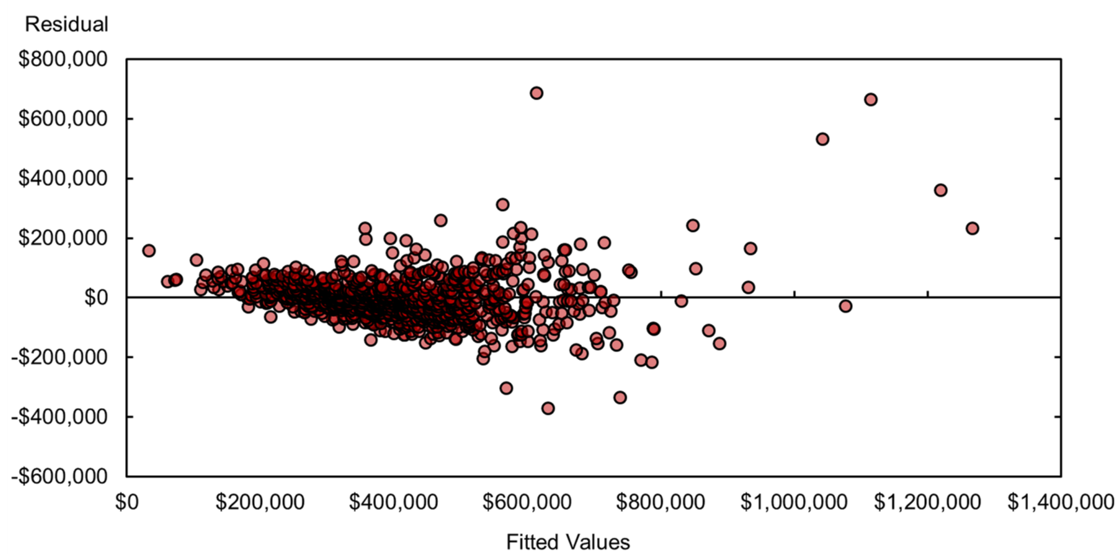


Chart 16: **Log-Linear Specification Hedonic Pricing Regression Residuals**

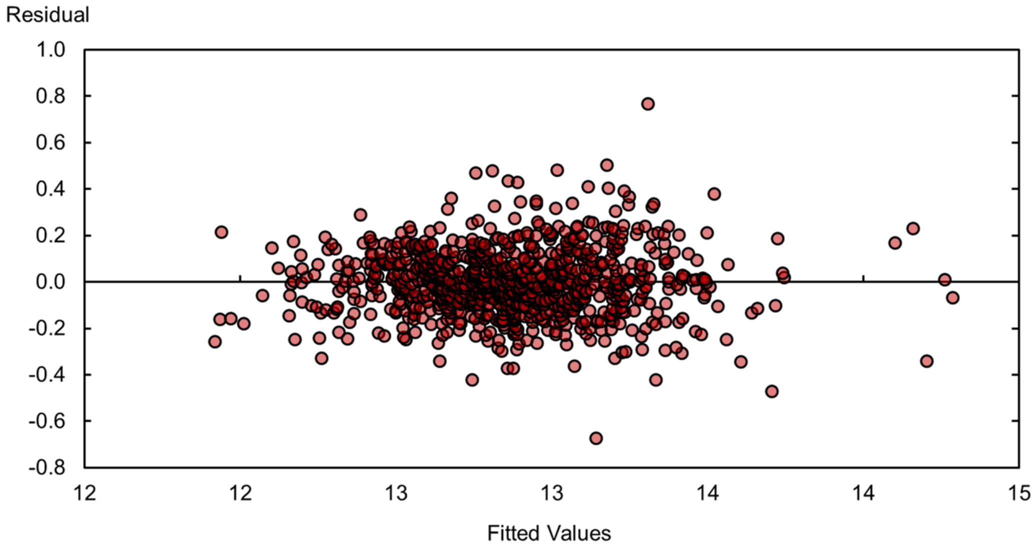
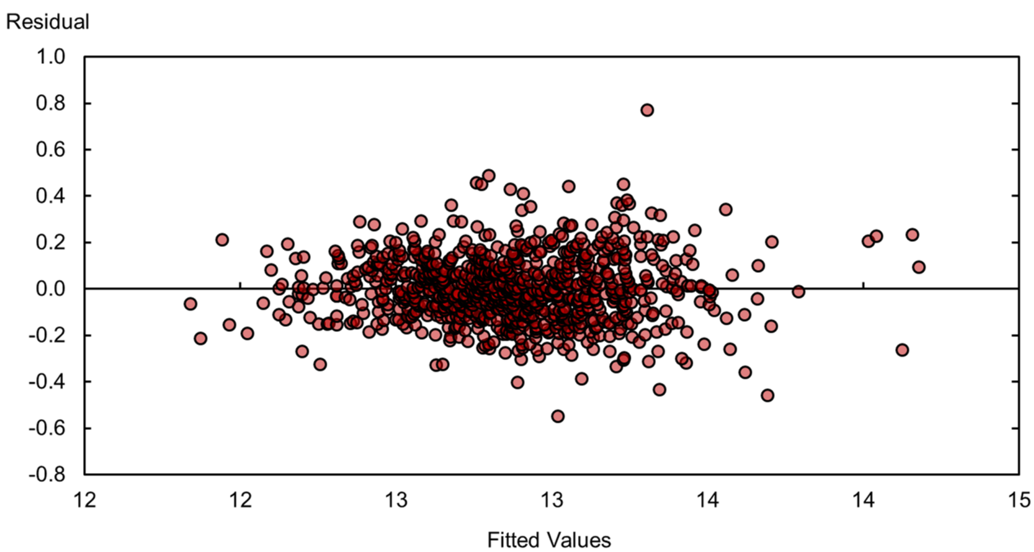


Chart 17: **Log-Log Specification Hedonic Pricing Regression Residuals**



Both the log-linear and log-log specifications have plausibly random residuals, and both fit almost equally well. These regressions also produce almost identical results. Therefore, the selected specification should not matter. The log-log specification was ultimately selected, as it allows for a slightly nicer interpretation. The following regression (**Table 4**) is the result of the log-log hedonic price regression (where none of the dummy variables were logged, and variables that were already percentages were also not transformed). An explanation of all independent variables can be found in appendix A (**Table 5**).

The Variance Inflation Factors (VIFs) were also reported for the hedonic regression, as collinearity can often be an issue. High collinearity between independent variables makes each seem less significant than they actually are but does not affect the model's predicted values or residuals. However, it can make coefficients appear unstable, as very high standard errors are reported for them. VIFs greater than 4 are often thought to be problematic. Only two variables, median household income and the per cent of people in a neighbourhood that are married, have VIFs that exceed this cut-off. As we are not particularly interested in either of these coefficients, the presence of collinearity can mostly be ignored.

Most variables are statistically significant at the five per cent level and have the expected sign. Interestingly, lot size appears to have no effect on sales price. However, this unusual result is likely an artifact of the data, where apartments have recorded lot sizes of zero square feet, and a small number of rural farms have very large lot sizes. The coefficients for bedrooms, square footage, bathrooms, garage, and fireplace are all in line with the meta-analysis discussed in Sirmans et al. (2005). In particular, the proxy for square footage is almost identical to the number reported in the Sirmans et al. paper, suggesting that the proxy selected seems to a good one.

Table 4

	(1) OLS	(2) Robust OLS	Variance Inflation Factor
Variable			
Logged_DaysontheMarket	0.0055	0.0055	1.08
	(0.0054)	(0.0057)	
ApartmentDummy	0.0979**	0.0979*	3.64
	(0.0430)	(0.0508)	
HouseDummy	0.0814***	0.0814***	1.70
	(0.0145)	(0.0134)	
LotArea	0.0000	0.0000*	1.07
	(0.0000)	(0.0000)	
TotalAGBedrooms	0.0536***	0.0536***	1.62
	(0.0080)	(0.0093)	
Basement_Bedrooms	0.0264***	0.0264***	1.91
	(0.0089)	(0.0082)	
TotalFullBaths	0.1215***	0.1215***	1.62
	(0.0086)	(0.0085)	
TotalHalfBaths	0.0601***	0.0601***	1.49
	(0.0105)	(0.0121)	
Logged_DistancefromCityHall	-0.0191***	-0.0191**	1.61
	(0.0071)	(0.0084)	
Garage_Code	0.1189***	0.1189***	1.51
	(0.0137)	(0.0151)	

Waterfront_code	0.4029***	0.4029***	1.18
	(0.0367)	(0.0524)	
PercentofHomesVacant	0.0049***	0.0049***	3.01
	(0.0009)	(0.0010)	
Logged_MedianHouseholdIncome	0.1556**	0.1556 **	16.19
	(0.0531)	(0.0557)	
MedianAge	0.0026**	0.0026**	2.74
	(0.0010)	(0.0011)	
Married	-0.0019	-0.0019	14.52
	(0.0014)	(0.0014)	
LessthanHighSchool	0.0004	0.0004	2.58
	(0.0014)	(0.0015)	
BacholorsDegreeorMore	0.0067***	0.0067***	3.57
	(0.0007)	(0.0008)	
Employed	-0.0014	-0.0014	3.52
	(0.0009)	(0.0010)	
Movedinthelastyear	0.0035***	0.0035***	2.64
	(0.0011)	(0.0011)	
Basementpresent	0.0647***	0.0647***	1.52
	(0.0230)	(0.0236)	
Propertytaxpercent	-0.6244***	-0.6244***	1.21
	(0.0400)	(0.0559)	
Tenantoccupied	-0.0326*	-0.0326*	1.22
	(0.0178)	(0.0166)	

Multipleunit	-0.0751**	-0.0751*	1.08
	(0.0326)	(0.0426)	
Vacant	-0.0097	-0.0097	1.18
	(0.0153)	(0.0169)	
Logged_CondoFees	-0.0169***	-0.0169***	3.17
	(0.0053)	(0.0064)	
Fireplacedummy	0.0661***	0.0661***	1.30
	(0.0115)	(0.0123)	
Logged_sq_ft_proxy	0.3562***	0.3562***	1.39
	(0.0229)	(0.0271)	
Constant	8.4744***	8.4744***	
	(0.5502)	(0.5705)	
n	867	867	
R^2	0.8366	0.8366	

The residuals from the hedonic regression represent all variation in sales price that the model was unable to explain. Assuming that the model is well specified, the residuals can be considered a measure of how much the final sales price differed from its "fundamental value." Assuming that the list price has no causal relationship with the sold price, the residuals should be unrelated to the list price.

This is a testable hypothesis. We now regress the log of the list price on the residuals (**Table 5**). The coefficient for list price is positive and highly statistically significant. This suggests that the list price is directly affecting how much a house sells for above its true value. Underpricing a house, it seems, significantly decreases how much it ultimately sells for holding its characteristics constant. These results are consistent with the dynamic search model findings. We then confine the regression to only examine houses that sold for above asking price, which is used as a proxy for a property experiencing a bidding war (**Table 6**). While the effect found is smaller, even conditional on a house experiencing a bidding war, undervalued houses sell for less than they would sell for if they were priced at fair market value. This is also consistent with the results from the dynamic search model.

While these results are consistent with the predictions of the dynamic search model, ideally we should also test the prediction that underpriced homes are more likely to sell. Unfortunately, while this very well may be true, we are unable to test for this. This is primarily because we do not have data on houses that did not sell and were taken off the market, leading to a possible selection effect. Furthermore, the hedonic regression found no statistically significant relationship between days on the market and final sold price. If relisting a property in order to reset the number of days on the market is common, it is reasonable to expect that no relationship will be found in the data.

Table 5

Variable	**(1) OLS**	**(2) Robust OLS**
Logged_ListPrice	0.1585*** (0.0123)	0.1585*** (0.0146)
n	867	867
R^2	0.1608	0.1608

Table 6		
Variable	**(1)** **OLS**	**(2)** **Robust OLS**
Logged_ListPrice	0.1087*** (0.02533)	0.1087*** (0.0286)
n	272	272
R^2	0.0639	0.0639

Other specifications for table 5 and 6 were also explored. Most notably, monthly dummies were tried in both the hedonic price specification and the residual regressions. In both cases, the dummies were not statistically significant, did not materially change the coefficients on the other independent variables, and did little to improve the fit of the model. Thus, at least in the short time period considered, no seasonal effects on prices could be detected.

Conclusion

This paper finds that, in the Kingston real estate market, underpricing a property leads to a lower overall sales price after controlling for a host of other variables. On average, underpricing a house by 1 per cent leads to a 0.16 per cent decrease in its sales price. In the case that an undervalued house experiences a bidding war, initially underpricing a house by 1 per cent leads to a 0.11 per cent decrease in its final sales price. These effects are quite large in magnitude. Underpricing the median property in the sample by 10 per cent would result in it selling for $6,100 less than pricing it at its true value (or $4,200 conditional on the house experiencing a bidding war).

There are a few possible explanations exist as to why a significant minority of sellers in Kingston underprice their homes. One possible explanation is that sellers use the list price to signal how motivated they are to sell their homes. While the MLS data does not allow us to test this hypothesis, it was explored using a dynamic search model. The model found that when sellers have different reservation prices, which could come about by some sellers being more motivated to sell than others, results in sellers with lower reservation prices receiving less for their otherwise identical house. However, they were also more likely to sell their homes. Therefore, this paper cannot rule out the explanation that underpricing a property is a method sellers employ to signal their level of motivation to sell.

Another possible explanation is that the Kingston real estate market has a principal-agent problem. Agents only receive a small percentage of the final sales price of a home, which often amounts to about 1.5 per cent of the final sales price. Thus, if a seller accepts an offer for $10,000 less than a house is worth, the agent only receives $150 less than she would if she sold it at its true value. If the agent believes that finding a buyer willing to pay more will take a significant amount of effort, she may advise the seller to accept the offer. Many academic papers suggest underpriced houses sell faster and are more likely to sell. Thus, this paper can also not rule out the theory that agents advise sellers to underprice a property and accept low offers order to make a quick, easy sale.

A final possible explanation is that real estate agents experience significant benefits from selling a house at or above list price. A Toronto real estate agent complained that Toronto real estate agents used their list-to-sales ratio to advertise their skills as an agent (Toronto Realty Blog, 2015). For example, a real estate agent may promise a seller top dollar by arguing that he has a 99.5 per cent list to sales ratio. A local Kingston real estate website (Remax, n.d.) advises sellers to only hire real estate agents with a history of selling a house for very close to its list price. If potential sellers are using sales to list price as a metric for choosing real estate agents, and real estate agents are advertising their high list to sales ratio or high rate of generating bidding wars in order to attract clients, agents may underprice homes in an effort to increase this ratio. In other words, Goodhart's law applies – while the list to sales price may be a metric that contains valuable information, once it becomes the target, it ceases to be a good measure of a real estate agent's ability to sell a home for top dollar.

Thus, this paper is not able to determine precisely why underpriced homes sell for less than their true value. A potential avenue for future research would be to explore the reason for this effect.

Bibliography

Albrecht, J., Gautier, P. A., Vroman, S. (2009). Directed Search in the Housing Market. IZA Discussion Paper.

Anglin, P., Arnott, R. (1991). Residential Real Estate Brokerage as a Principal-Agent Problem. The Journal of Real Estate Finance and Economics, 99-125.

Ashenfelter, O., Genesove, D. (1992). Testing for Price Anomalies in Real-Estate Auctions. American Economic Review, 501-505.

Beracha, E., Seiler, M. (2014). The Effect of Listing Price Strategy on Transaction Selling Prices. Journal of Real Estate Finance and Economics, 237-255.

Bucchianeri, G. W., Minson, J. A. (2013). A homeowner's dilemma: Anchoring in residential real estate transactions. Journal of Economic Behavior Organization, 76-92.

Butler, R. V. (1982). The specification of hedonic indexes for urban housing. Land, 94-108.

Cassel, E., Mendelsohn, R. (1985). The choice of functional forms for hedonic price equations: Comment. Journal of Urban Economics, 135-142.

CF, S., GK, T., J, D. (1997). Residential development, risk, and land prices. J Reg Sci, 613–628.

Chau, K., Chin, T. L. (2003). A Critical Review of Literature on the Hedonic Price Model. International Journal for Housing Science and Its Applications, 145-165.

City of Kingston. (2018). City of Kingston - 2018 Tax Rates. Retrieved from City of Kingston: https://www.cityofkingston.ca/documents/10180/57933/2018+Tax+Rate+Schedule/ 8606-4ba4-a999-509de328b1ba

Competition Bureau. (2018, November 13). The courts have been loud and clear in TREB case: Time to move forward. Retrieved from Competition Bureau: https://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/04402.html

Cotteleer, G., Gardebroek, C., Luijt, J. (2008). Bargaining and market power in a GIS-based hedonic pricing model of the agricultural land market. Land Economics, 573-592.

Cropper, M. L., Deck, L. B., McConnell, K. E. (1988). On the Choice of Functional Form for Hedonic Price Functions. The Review of Economics and Statistics, 668-675.

Cubbin, J. (1974). Price, quality, and selling time in the housing market. Applied Economics, 171-187.

Diewert, E. (2003). Hedonic Regressions. A Consumer Theory Approach. In R. C. Shapiro, National Bureau of Economic Research - Scanner Data and Price Indexes (pp. 317-348). University of Chicago Press.

Garrod, G., Willis, K. (1992). The amenity value of woodland in Great Britain: A comparison of economic estimates. Environmental Resource Economics, 415-434.

Geltner, D., Kluger, B. D., Miller, N. G. (1991). Optimal Price and Selling Effort from the Perspective of the Broker and Seller. Real Estate Economics, 1-24.

Genesove, D., Mayer, C. (1997). Equity and Time to Sale in the Real Estate Market. American Economic Review, 255-269. Glower, M., Haurin, D. R., Hendershott, P. H. (1998). Selling Time and Selling Price: The Influence of Seller Motivation. Real Estate Economics, 719-740.

Kalinowski, T. (2018, September 11). Real estate companies threatened by TREB after posting GTA sales data. Retrieved from The Star: https://www.thestar.com/business/real$_{est}$ $estate - company - threatened - by - toronto - board - after - posting - sales - data.html$

Kuminoff, N., Parmeter, C. F., Pope, J. C. (2010). Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? Journal of Environmental Economics and Management, 145-160.

Levitt, S. D., Syverson, C. (2008). Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions. Review of Economics and Statistics, 599-611.

Linneman, P. (1980). Some empirical results on the nature of the hedonic price function for the urban housing market. Journal of Urban Economics, 47-68.

Lisi, G., Iacobini, M. (2012). Measuring the Housing Price Dispersion in Italy. Economics Bulletin, 811-822.

Maury, T.-P., Tripier, F. (2014). Search strategies on the housing market and their implications on price dispersion. Journal of Housing Economics, 55-80.

Merlo, A., Ortalo-Magne, F. (2004). Bargaining over residential real estate: evidence from England. Journal of Urban Economics, 192-216.

Miller, N., Sklarz, M. (1987). Pricing Strategies and Residential Property Selling Prices. Journal of Real Estate Research, 31-40.

Rasmussen, D. W., Zuehlke, T. W. (2006). On the choice of functional form for hedonic price functions. Applied Economics, 431-438.

Remax. (n.d.). Preparing to Sell. Retrieved from Remax Kingston: http://www.hollyandluca.com/

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. The Journal of Political Economy, 34-55.

Rutherford, R., Yavas, A. (2012). Discount Brokerage in Residential Real Estate Markets. Real Estate Economics, 508-535.

Rutherford, R., Springer, T., Yavas, A. (2005). Conflicts between principals and agents: evidence from residential brokerage. Journal of Financial Economics, 627-665.

Sirmans, G. S., MacDonald, L., Macpherson, D. A., Zietz, E. N. (2006). The Value of Housing Characteristics: A Meta Analysis. The Journal of Real Estate Finance and Economics, 215-240.

Statistics Canada. (2018, December 14). National balance sheet and financial flow accounts, third quarter 2018. Retrieved from statcan: https://www150.statcan.gc.ca/n1/daily-quotidien/181214/dq181214a-eng.htm

Statistics Canada. (2019, July 31). Canada Mortgage and Housing Corporation, vacancy rates, apartment structures of six units and over, privately initiated in census metropolitan areas. Retrieved from statcan: https://www150.statcan.gc.ca/t1/tbl1/en/tv.actio

Toronto Realty Blog. (2015, January 28). Does "Sale-To-List" Ratio Matter? Retrieved from Toronto Realty Blog: https://torontorealtyblog.com/blog/does-list-to-sale-ratio-matter/

Tucker, C., Zhang, J., Zhu, T. (2013). Days on market and home sales. The RAND Journal of Economics.

Zolo. (2019, July 29). Retrieved from Toronto Real Estate Trends - Housing Market Report for July 2019: https://www.zolo.ca/toronto-real-estate/trends

Appendices

Appendix A

Table 5: Description of hedonic pricing variables

Logged_DaysontheMarket	Number of days the property was on the market before it sold (natural log)
ApartmentDummy	Equals 1 if the property is listed as being an apartment (dummy)
HouseDummy	Equals 1 if the property is listed as being an house (dummy)
LotArea	Square footage of the lot size (0 for an apartment)
TotalAGBedrooms	Number of bedrooms above ground level
Basement_Bedrooms	Number of bedrooms in the basement
TotalFullBaths	Number of full bathrooms
TotalHalfBaths	Number of half bathrooms
Logged_DistancefromCityHall	Distance in km from city hall (natural log)
Garage_Code	Equals 1 if the property has a garage (dummy)
Waterfront_code	Equals 1 if the property has a waterfront view (dummy)
PercentofHomesVacant	Percent of houses vacant in the neighbourhood (2016 census, %)

Logged_MedianHouseholdIncome	Median income for the neighbourhood (2016 census, $)
MedianAge	Median age of the neighbourhood (2016 census, age)
Married	Percent of neighbourhood residents that are married (2016 census , %)
LessthanHighSchool	Percent of neighbourhood residents that have less than a high school diploma (2016 census , %)
BacholorsDegreeorMore	Percent of neighbourhood residents that have a bachelor's degree or more (2016 census, %)
Employed	Percent of neighbourhood residents that are employed (2016 census , %)
Movedinthelastyear	Percent of neighbourhood residents that moved in the last year (2016 census , %)
Basementpresent	Equals 1 if the property is listed has a basement (dummy)
Propertytaxpercentfix	Percent of property tax paid as a per cent based on sales price instead of assessed value (%)
Tenantoccupied	Equals 1 if the property is listed as being occupied by a tenant (dummy)
Multipleunit	Equals 1 if the property is listed as being occupied by a tenant and the owner (dummy)

Vacant	Equals 1 if the property is listed as being vacant (dummy)
Logged_CondoFees	Monthly condo fees amount - zero if no condo fees imposed (natural log)
Fireplacedummy	Equals 1 if the property is listed as having a fireplace (dummy)
Logged_sq_ft_proxy	Number of square feet - proxy (natural log)

Appendix B

The two pages that follow are an example of a Kingston MLS PDF. Almost all housing characteristics for the project were derived from these PDFs. Some potentially identifying information has been deliberately redacted.

RESIDENTIAL

List Price $399,900
Status SOLD
MLS # K19001719 **(Orig #** **)**
Type Detached
SubType Residential
Land Type Deeded Land
Title Freehold
Sale Type
Style 2 Storey

DOM

Major Area	Kingston	**Waterfront Y/N**	
District	Kingston	**Waterfront Name**	
Sub District	[redacted]	**Waterfrontage (ft)**	
Address	[redacted]	**Shoreline Rd Allow**	
City/Twn/Muni	Kingston , ON	**Waterfront Exposure**	
Postal Code	[redacted]	**WF Features**	
County	Kingston	**WF Features Other**	
Seller Name 1	[redacted]	**Shoreline**	
Seller Name 2		**Lot Frontage/Depth**	**Irregular?**
Add'l Sellers		**Acreage**	
PIN	[redacted]	**Occupancy**	Vacant
Roll #	[redacted]	**Sign Y/N**	
Acreage	**Property Size** 0.0 - 0.49 Acres	**Possession**	
Zoning	Residential	**Annual Taxes $**	$3,526.00 **Tax Year** 2018
Age (Building)	11-20 Years	**Association Fee**	
Fronting On	West	**Land Lease Fee**	**Road Maint Fee**
		Rental Income	

Legal Description PT LT 60, PL13M34, PT26 13R16274
Directions [redacted]
Realtor Remarks 24 hours irrevocable
Public Remarks A wonderful 3+1 bedroom, 3.5 bathroom home on the [redacted] Entering the main level you will find a separate dining room, eat-in kitchen with sliding doors to the deck, and a large living room featuring cathedral ceilings and plenty of natural light. This level also has a 2 pc bathroom, access to the garage, and is carpet free with slate and hardwood flooring. Upstairs you will find 3 bedrooms, the main bathroom and an ensuite for the master. The lower level is fully finished with a family room, extra bedroom, 3 pc bath, laundry and utility room. [redacted] the back yard features a deck, a gazebo and backs on to a treed area from [redacted]. Recent upgrades include the roof, fridge, stove, washer and dryer all in 2015. Available for occupancy anytime after May 6, 2019.

CONDOMINIUM

Condo/Lot Fee $/Month
Condo # Parking Spaces **Parking Space #**
Condo Locker # **Special Assmt**
Prop Mgmt Co
Prop Mgmt Phone
Prkg Space Ownership

EXPENSES

Annual Hydro Costs $ **Annual Heat Cost $**
Annual Water/Sewer $ **Annual Lot Rental $**
Annual Rental Equip $

GENERAL INFORMATION

Total AG SqFt **Approx SqFt Range**
Heating Source Natural Gas
Hot Water Heating Natural Gas
Heating Type Forced Air
Fireplace Type
Stove Type
Access Municipal Road
Other Access Info
Documents Avail

MLS # [redacted]

Services Avail Cable, Garbage PickUp, High Speed Internet, Natural Gas, Cell Service
Water/Well Municipal Water
Sewer/Septic Sewer
Basement Full Basement
Bsmt Finish Fully Finished
Foundation Block
Vermiculite Insul
Ext Finish Brick, Vinyl Siding
Roof Type Asphalt Shingle
Amperage **Alt Power**
Parking Spc **Garage Y/N** Yes **Garage Features** 1 Car, Attached, Inside Entry
Driveway Size Single
Driveway Det. Asphalt
Features Int Ceilings-greater than 8', Central Air, Ensuite, Rec Room
Rental Equip Water Heater
Features Ext Deck, Fenced Yard, Gazebo
Restrictions
Site Influences
Fixtures Excluded NONE
Chattel Included
CHATTELS INCLUDED (K) Dishwasher, Dryer, Gazebo, Refrigerator, Stove, Washer

ROOMS AND THEIR SIZES

 # Bedrooms AG 3 **# Total Bedrooms** 4
 # Full Baths 3 **# Half Baths** 1 **Total Baths** 4

ROOM SIZE	LEVEL	TYPE	FLOORING	ROOM SIZE	LEVEL	TYPE	FLOORING
11'6" X 11'4"	Main Floor	Dining Room		10'8" X 9'9"	2nd Floor	Bedroom	
17'6" X 9'9"	Main Floor	Kitchen		17'7" X 12'3"	Basement	Rec Room	
18'8" X 13'3"	Main Floor	Living Room		11'11" X 8'	Basement	Bedroom	
	Main Floor	2 Pc Bathroom		7'6" X 6'10"	Basement	Laundry	
17' X 11'4"	2nd Floor	Master Bedroom			Basement	Utility Room	
	2nd Floor	4 Pc Ensuite			Basement	3 Pc Bathroom	
	2nd Floor	4 Pc Bathroom					
11'1" X 10'10"	2nd Floor	Bedroom					

LISTING AGENT/OFFICE

List Agent 1A [redacted]
List Office 1A [redacted]
List Agent 1B
List Office 1B
List Agent 2
List Office 2
Co-Op Broker Comm $ $0.00 **Co-Op Broker Comm %** 2.00

STATUS & DATES

List Date 4/1/2019 **Expiry Date** **Contact Permitted** No
Publish to Realtor.ca **Publish Add to Pub Sites** Yes **Publish to Public Sites** Yes
V Tour Public Access **Virtual Tour** Virtual Tour
Virtual Tour 2 **Virtual Tour 3**
DDF Include Y/N Yes **DDF Linkback URL**
Sales Brochure URL **Audio Input URL**

SELLING AGENT/OFFICE

HOW SOLD SOLD **Sold Price** $390,000 **Firm (Sold) Date** 5/14/2019 **Closing Date** 7/12/2019
Sell Agent 1A [redacted] **Sell Off 1A** [redacted]
Sell Agent 1B **Sell Off 1B**
Sell Agent 2 **Sell Off 2**

LISTING COURTESY OF:

[redacted]

[redacted]