



UNIVERSITÉ DU QUÉBEC EN OUTAOUAIS

DÉPARTEMENT DES SCIENCES ADMINISTRATIVES

WORKING PAPER NO. 1

Heterogeneous Amenity Capitalization in the Quebec Residential Market: A Spatial Quantile Regression Approach

Simon-Pierre Boucher

Département des sciences administratives

Université du Québec en Outaouais

simon-pierre.boucher@uqo.ca

Gatineau – Pavillon Alexandre-Taché

283, boulevard Alexandre-Taché

Gatineau, Québec, Canada J9A 1L8

May 2026

Version 1.0

Abstract

Standard hedonic pricing models estimated by ordinary least squares impose constant implicit prices for housing attributes across the conditional price distribution—an assumption that is difficult to justify when buyers at different price segments have different preferences, constraints, and substitution possibilities. This paper examines whether neighborhood amenity capitalization varies across the price distribution in the Quebec residential market. Using 42,481 property listings from Centris spanning four market segments—single-family houses, condominiums, multiplexes, and vacant land—we estimate quantile regression models at $\tau \in \{0.10, 0.25, 0.50, 0.75, 0.90\}$ and augment them with an instrumented spatial lag in a two-stage spatial quantile regression (2SQR) framework. Inter-quantile Wald tests reject the null of coefficient homogeneity (omnibus $\chi^2(7) = 35.74$, $p < 0.001$ for houses). The results suggest that hospital proximity shifts from a significant negative association (-0.082) at the 10th percentile to a positive (insignificant) association at the 90th percentile; park proximity declines from $+0.087$ ($p < 0.001$) to -0.044 (insignificant); and transit proximity declines from $+0.074$ to -0.008 across the distribution. The spatial lag parameter ρ is positive and significant across all segments, with Moran's I on OLS residuals ranging from 0.44 to 0.55 and first-stage F -statistics exceeding 100 in all segments. Because the analysis relies on listing prices and cross-sectional variation, the estimated coefficients are best interpreted as quantile-specific capitalization gradients rather than causal effects. To our knowledge, this is the first multi-segment spatial quantile analysis of amenity capitalization in the Quebec housing market.

Keywords: hedonic pricing, spatial quantile regression, amenity capitalization, housing market segmentation, implicit prices, Quebec real estate, Centris, bootstrap inference

JEL Classification: R31, C21, C31, Q51

1. Introduction

The hedonic pricing model, formalized by [Rosen \(1974\)](#), remains the workhorse framework for estimating the implicit prices of housing attributes. By regressing property prices on structural characteristics, locational features, and neighborhood amenities, researchers can decompose market values into the marginal willingness to pay (WTP) for each attribute. This approach has generated an extensive empirical literature across international housing markets ([Malpezzi, 2003](#); [Sirmans et al., 2005](#); [Chin & Chau, 2003](#)).

However, standard hedonic models estimated by ordinary least squares (OLS) impose a critical assumption: that implicit prices are constant across the conditional price distribution. A household purchasing a \$150,000 property is assumed to value an additional bathroom or park proximity identically to a household purchasing a \$1,500,000 property. This homogeneity assumption has been increasingly questioned on both theoretical and empirical grounds ([Zietz et al., 2008](#); [Liao & Wang, 2012](#); [Mak et al., 2010](#)). From a theoretical perspective, [Rosen \(1974\)](#) himself acknowledged that the hedonic price function reflects a market equilibrium between heterogeneous buyers and sellers, and there is no reason to expect the marginal valuation of attributes to be invariant across market participants. From an empirical standpoint, housing markets are characterized by substantial preference heterogeneity: budget-constrained households prioritize functional accessibility (transit, highways), while luxury-market buyers emphasize prestige, privacy, and environmental quality ([Waddell et al., 1993](#); [Bayer et al., 2007](#)).

Two methodological developments address this limitation. First, quantile regression ([Koenker & Bassett, 1978](#); [Koenker, 2005](#)) allows coefficients to vary across the conditional distribution, revealing how implicit prices differ between affordable and luxury properties. Unlike mean regression, quantile regression characterizes the entire conditional distribution of the response variable, providing a more complete picture of the relationship between housing prices and their determinants ([Hao & Naiman, 2007](#)). Second, spatial econometric methods ([LeSage & Pace, 2009](#); [Anselin, 1988](#)) account for the well-documented spatial dependence in housing markets, where neighboring property prices and characteristics influence valuations through spillover effects. The combination of these two approaches—spatial quantile regression—remains relatively rare in the literature, despite its theoretical appeal ([Liao & Wang, 2012](#); [Su & Yang, 2007](#)).

The Canadian housing market presents a particularly interesting case for studying amenity capitalization heterogeneity. Canada experienced one of the most pronounced housing price escalations among OECD countries between 2000 and 2022, with the composite benchmark price rising over 300% nationally ([CMHC, 2023](#)). Within Canada, the Quebec market exhibits distinctive characteristics: a bilingual regulatory environment, a unique civil law tradition affecting property transactions, and significant price variation between the Montreal metropolitan area and

regional markets (Statistics Canada, 2023). The province's housing stock spans from dense urban condominiums in downtown Montreal to rural single-family homes and agricultural land parcels, creating natural variation for studying how amenity capitalization differs across market segments and price levels.

Despite these features, the Canadian housing literature—and the Quebec market in particular—remains largely confined to mean-based hedonic estimates. Recent Quebec-focused studies have examined social housing impacts (Dubé et al., 2023), park premiums (Dubé et al., 2025), supervised consumption site effects (Perreault et al., 2024), and rail transit capitalization (Dubé et al., 2011), but none have investigated distributional heterogeneity in amenity capitalization using spatial quantile methods. The broader Canadian spatial hedonic literature includes studies of Toronto (Des Rosiers et al., 2011), Vancouver (Pace & LeSage, 2004), and pan-Canadian analyses (Bourassa et al., 2010), but quantile regression applications remain scarce.

This paper makes four contributions to the literature. First, we provide the first application of spatial quantile regression to the Quebec residential market using a comprehensive dataset of 42,481 Centris listings. Second, we extend the analysis beyond single-family homes to four distinct market segments—houses, condominiums, multiplexes, and vacant land—enabling cross-segment comparison of amenity capitalization patterns. Third, we employ a two-stage spatial quantile regression (2SQR) framework with bootstrap inference, combining the distributional flexibility of quantile regression with explicit modeling of spatial spillovers. Fourth, we implement a comprehensive robustness framework including weak instrument diagnostics, variance inflation factor (VIF) analysis, sensitivity tests on spatial weights parameterization, and omnibus inter-quantile heterogeneity tests, providing confidence in the stability of our findings.

Our results reveal substantial heterogeneity. For single-family homes, the omnibus inter-quantile test strongly rejects coefficient homogeneity ($\chi^2(7) = 35.74$, $p < 0.001$), with hospital proximity, park proximity, and transit proximity exhibiting statistically significant differences between the 10th and 90th percentiles. Notably, hospital proximity shifts from a significant discount of 8.2% for lower-priced homes to a positive (but insignificant) association at the upper tail; park proximity declines from a significant premium of 8.7% to an insignificant negative coefficient. The spatial lag parameter ranges from 0.04 to 0.23 across segments at the median, with first-stage F -statistics exceeding 100 in all segments, well above the Stock & Yogo (2002) threshold for instrument relevance. The OLS model with corrected distance measures explains 68.5% of house price variation (adjusted $R^2 = 0.685$). These findings have direct implications for urban planning, property taxation, and the interpretation of amenity valuation studies that rely on conditional-mean estimates.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature on hedonic pricing, spatial econometrics, quantile regression, and housing market segmentation. Section 3 presents the econometric framework, including the 2SQR estimator, bootstrap inference,

and diagnostic tests. Section 4 describes the data source, variable construction, and sample characteristics. Section 5 reports the empirical results across all four market segments. Section 6 presents the robustness checks. Section 7 discusses theoretical and policy implications, and Section 8 concludes with directions for future research.

2. Literature Review

2.1. The Hedonic Pricing Framework

The theoretical foundation of hedonic pricing rests on Rosen (1974), who showed that in a market with differentiated products, the equilibrium price function reveals marginal valuations of product attributes. In housing markets, the hedonic price function $P = f(S, L, N)$ relates property prices to structural characteristics S , locational attributes L , and neighborhood amenities N . The partial derivative $\partial P / \partial z_k$ yields the implicit (hedonic) price of attribute z_k , interpreted as the market's marginal willingness to pay.

The log-linear specification $\ln P = \alpha + \beta' \mathbf{X} + \varepsilon$ has become the dominant functional form (Sirmans et al., 2005; Malpezzi, 2003; Halvorsen & Pollakowski, 1981). In this specification, coefficients on continuous variables represent semi-elasticities, while coefficients on dummy variables approximate percentage effects via $(\exp(\beta) - 1) \times 100$. Cropper et al. (1988) compare alternative functional forms (linear, semi-log, double-log, Box-Cox) using Monte Carlo simulations and find that the log-linear form performs well under omitted variable bias, a pervasive concern in hedonic studies. Cassel & Mendelsohn (1985) further demonstrate that the semi-log form yields more interpretable and stable coefficient estimates than flexible alternatives when the true functional form is unknown.

A recent methodological reflection by Herath & Maier (2025) emphasizes persistent challenges: functional form misspecification, omitted variable bias, spatial dependence, and the assumption of a single market equilibrium. Parmeter & Pope (2013) provide a comprehensive review of applied hedonic methods, noting that the choice of functional form, variable selection, and spatial structure can substantially affect estimated implicit prices. These challenges motivate the multi-method approach adopted in this paper.

The identification problem in hedonic models deserves particular attention. As Ekeland et al. (2004) and Bajari & Benkard (2005) demonstrate, identifying the demand function for housing attributes requires either functional form restrictions or instruments. In a single cross-section, the hedonic price function reflects the equilibrium between supply and demand; the implicit prices we estimate are therefore market-clearing marginal values rather than structural demand parameters. We acknowledge this limitation while noting that for policy analysis (e.g., property taxation, cost-benefit analysis of amenity investments), the equilibrium implicit prices are the relevant quantities.

2.2. Spatial Hedonic Models

Housing prices exhibit strong spatial dependence: properties in the same neighborhood share unobserved characteristics (school quality, crime rates, aesthetic appeal) that generate correlated residuals. Ignoring spatial autocorrelation produces biased and inefficient OLS estimates (LeSage & Pace, 2009; Anselin, 1988; Can, 1992).

The spatial autoregressive (SAR) model augments the hedonic equation with a spatial lag of the dependent variable:

$$\ln \mathbf{P} = \rho \mathbf{W} \ln \mathbf{P} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (1)$$

where $\mathbf{W}$ is a spatial weights matrix and ρ captures the degree of spatial spillover. The spatial error model (SEM) places spatial structure in the disturbance term $\boldsymbol{\varepsilon} = \lambda \mathbf{W}\boldsymbol{\varepsilon} + \mathbf{u}$, while the spatial Durbin model (SDM) nests both by including spatially lagged explanatory variables $\mathbf{WX}$ (LeSage & Pace, 2009; Elhorst, 2010). Anselin (2010) provides a comprehensive review of thirty years of spatial econometrics, documenting the evolution from simple spatial lag and error models to more flexible specifications.

The interpretation of spatial models requires attention to the distinction between direct and indirect effects. In the SAR model, the marginal effect of a change in x_{ik} on y_i includes not only the direct effect β_k but also feedback effects through the spatial multiplier $(I - \rho W)^{-1}$. LeSage & Pace (2009) formalize this decomposition, showing that the average total effect of a unit change in x_k across all observations is $n^{-1}\text{tr}[(I - \rho W)^{-1}]\beta_k$, which can be approximated as $\beta_k/(1 - \rho)$ for global effects.

The specification of the spatial weights matrix $\mathbf{W}$ is a critical modeling choice. Common approaches include contiguity-based weights, distance bands, and K -nearest neighbors (KNN) (Getis, 2004). Stakhovych & Bijmolt (2009) demonstrate that misspecification of $\mathbf{W}$ can bias spatial parameter estimates, though the impact on non-spatial coefficients is often modest. KNN weights offer the practical advantage of avoiding “island” observations (properties with no neighbors under distance-band specifications) in datasets with irregular spatial density (Pace & LeSage, 2004).

Recent advances combine spatial structure with machine learning. Gyger et al. (2025) propose a hybrid spatial Gaussian process with tree boosting, tested on 1.5 million German rental listings. Chen et al. (2025) apply GeoShapley to decompose spatial effects in Beijing housing prices. For the Canadian context, Bourassa et al. (2010) document significant spatial dependence in Canadian housing markets, and Dubé et al. (2018) confirm this for the Quebec market specifically. Dubé et al. (2011) estimate a spatial hedonic model of rail transit capitalization in Montreal, finding that the spatial lag absorbs approximately 20% of the transit premium estimated by OLS.

2.3. Quantile Regression in Housing Markets

Quantile regression (Koenker & Bassett, 1978; Koenker, 2005) estimates the conditional quantile function:

$$Q_{\tau}(\ln P|\mathbf{X}) = \mathbf{X}\boldsymbol{\beta}(\tau), \quad \tau \in (0, 1) \quad (2)$$

where $\boldsymbol{\beta}(\tau)$ varies with the quantile τ . This reveals how the implicit price of each attribute differs across the price distribution—a dimension invisible to OLS, which estimates only $E[\ln P|\mathbf{X}]$.

The econometric advantages of quantile regression over OLS are well-documented. First, quantile regression estimators are consistent under heteroskedasticity and do not require distributional assumptions on the error term (Koenker, 2005). Second, they are more robust to outliers than OLS, since the check function penalizes deviations asymmetrically rather than quadratically (Buchinsky, 1998). Third, quantile regression reveals the full structure of the conditional distribution, enabling researchers to test whether the effect of a covariate is heterogeneous across the response distribution (Hao & Naiman, 2007).

Zietz et al. (2008) provide an influential early application to U.S. housing data, finding significant heterogeneity in housing attribute coefficients across quantiles. They document that structural variables (bedrooms, bathrooms, floor area) tend to have larger effects at upper quantiles, while locational variables show more complex patterns. Liao & Wang (2012) extend this to a spatial quantile framework using Shanghai housing data, documenting that implicit prices of structural and neighborhood attributes vary substantially between cheap and expensive properties. Mak et al. (2010) confirm these patterns for the Hong Kong market, finding that floor area effects are three times larger at the 90th percentile than at the 10th percentile. McMillen (2008) applies quantile regression to land values in Chicago, finding that the distance gradient from the CBD varies substantially across quantiles. Farmer & Lipscomb (2010) use quantile regression to show that the premium for school quality is exaggerated at higher price quantiles in the U.K. housing market.

In the Canadian context, quantile regression applications to housing markets remain rare. Des Rosiers et al. (2014) apply quantile regression to assess the impact of environmental contamination on housing values in Montreal, finding that contamination effects are amplified for lower-priced properties. Wen et al. (2017) estimate quantile regression models for Toronto condominium prices, documenting heterogeneous effects of building age and floor level across the price distribution. These Canadian studies do not, however, combine quantile regression with spatial modeling.

The combination of spatial structure and quantile estimation remains methodologically challenging. Kim & Muller (2004) propose a two-stage instrumental variables approach: the spatial lag $\mathbf{W} \ln \mathbf{P}$ is instrumented with spatially lagged exogenous variables $\mathbf{W}\mathbf{X}$ in a first stage, then the predicted spatial lag is included in a second-stage quantile regression. This two-stage spatial quantile regression (2SQR) avoids the computational intractability of jointly estimating spatial and quantile

parameters. [Su & Yang \(2007\)](#) propose a profile likelihood approach to spatial quantile regression, while [Kostov \(2009\)](#) introduces a penalized quantile regression method with spatial smoothing. The 2SQR framework has been applied to airport noise effects ([Rehman & Pennington-Cross, 2025](#)), hospital accessibility ([Diao & Ferreira, 2023](#)), and urban green space valuation ([Chen et al., 2020](#)), but not to multi-segment Canadian housing data.

[Salvati et al. \(2023\)](#) propose a spatial semiparametric M-quantile regression framework for hedonic pricing, offering robustness to outliers and distributional misspecification. Their approach accounts for nonlinear environmental effects while maintaining spatial trend modeling. [Chernozhukov & Hansen \(2006\)](#) develop instrumental variable quantile regression methods with formal weak instrument diagnostics, which inform our testing strategy.

2.4. Housing Market Segmentation

The assumption that a single hedonic price function governs the entire market has been challenged by the housing submarket literature. [Bourassa et al. \(2003\)](#) demonstrate that segmented hedonic models outperform pooled models for property valuation, using data from Auckland, New Zealand. [Goodman & Thibodeau \(1998\)](#) provides theoretical foundations for market segmentation based on supply rigidities and demand heterogeneity. [Rey-Blanco et al. \(2024\)](#) use machine learning to identify spatial market segments endogenously in Madrid, finding that data-driven boundaries outperform administrative ones.

Cross-segment analysis—comparing hedonic estimates across property types such as single-family homes, condominiums, and multiplexes—has received less attention. Property types differ fundamentally in their attribute bundles: condominiums lack land, multiplexes generate rental income, and vacant land has no building characteristics. These structural differences imply that amenity capitalization should vary across segments, yet most hedonic studies analyze a single property type ([Sirmans et al., 2005](#)). [Palm \(1978\)](#) is among the earliest to argue that hedonic models should be estimated separately by housing submarket, while [Watkins \(2001\)](#) proposes formal tests for market segmentation based on Chow statistics. [Pryce & Oates \(2008\)](#) emphasizes that cross-segment analysis is essential for understanding how different buyer populations value neighborhood amenities.

2.5. Amenity Capitalization

The capitalization of neighborhood amenities into property values has been extensively studied. Parks generate premiums of 3–10% for proximate properties ([Crompton, 2001](#); [Anderson & West, 2006](#); [Troy & Grove, 2008](#)), though effects diminish with distance and may vary with park type and quality. School quality premiums are well-documented ([Black, 1999](#); [Gibbons & Machin,](#)

2008), though proximity to schools can generate disamenity effects from noise and traffic congestion. Transit accessibility premiums vary widely depending on transit mode and neighborhood context (Debrezion et al., 2007; Mohammad et al., 2013), with heavy rail stations typically commanding larger premiums than bus stops. Highway proximity produces competing effects: accessibility benefits versus noise, pollution, and visual disamenity (Theebe, 2004; Nelson, 2004). Hospital proximity effects are less studied but potentially complex, combining accessibility benefits with institutional disamenity effects (Diao & Ferreira, 2023).

For the Quebec context, Dubé et al. (2025) find additional price premiums for houses exposed to urban parks in Québec City, beyond simple proximity effects. Dubé et al. (2023) document asymmetric impacts of social housing on nearby property values, with NIMBY effects dominating in upper-income neighborhoods. Dubé et al. (2011) estimate a spatial hedonic model showing that proximity to Montreal's commuter rail network generates a 4–6% price premium. However, no published study has examined whether these amenity effects are heterogeneous across the price distribution in the Quebec market.

The literature on golf course proximity effects on property values is relatively thin compared to parks and transit. Do & Grudnitski (1995) find that golf course proximity generates a 7.6% premium, while Grudnitski & Do (2002) estimate premiums of 4–8% depending on lot position relative to fairways. These effects may vary with the socioeconomic profile of the neighborhood and the quality of the golf facility.

2.6. Bootstrap Inference in Spatial Models

Standard asymptotic inference for quantile regression can be unreliable when the data exhibit spatial dependence, heteroskedasticity, or clustering (Buchinsky, 1995; Hahn, 1995). The bootstrap provides a flexible alternative that avoids the need to estimate the density of the error term at zero—a requirement of the asymptotic sandwich variance estimator for quantile regression (Koenker, 2005).

In spatial contexts, the standard nonparametric (paired) bootstrap resamples observations with replacement, which breaks the spatial structure of the data. Lahiri (2003) discusses block bootstrap methods that preserve local spatial dependence by resampling contiguous blocks of observations. Jin & Su (2001) proposes a residual-based spatial bootstrap, while Freedman (1981) establishes the theoretical foundations for bootstrapping regression models. For instrumental variable quantile regression, Chernozhukov & Hansen (2006) develop formal inference methods, but these are computationally prohibitive for large datasets with spatial structure.

We employ the nonparametric paired bootstrap with $B = 200$ replications, acknowledging that this approach may not fully account for spatial dependence in the resampling scheme. Our bootstrap standard errors should therefore be interpreted as potentially optimistic; we discuss sensitivity to the number of replications and the implications of ignoring spatial structure in the bootstrap in

Section 6.

2.7. Gaps in the Literature

Four gaps motivate this study:

1. **No spatial quantile hedonic analysis for Quebec.** Despite extensive hedonic work by the Université Laval group (Dubé et al., 2023, 2025, 2011) and emerging studies in other Canadian markets (Des Rosiers et al., 2014; Wen et al., 2017), distributional heterogeneity in amenity capitalization has not been investigated using spatial quantile methods in Quebec.
2. **No multi-segment comparison.** Cross-property-type analysis of amenity effects is rare in any market (Sirmans et al., 2005), and absent for Quebec. Comparing houses, condos, multiplexes, and land provides unique insight into how buyer heterogeneity drives amenity capitalization.
3. **Under-exploitation of Centris data.** The Centris platform provides the most comprehensive Quebec listing database, yet published academic studies using it remain few. The breadth and detail of Centris listings enable analyses that are infeasible with smaller, survey-based datasets.
4. **Insufficient robustness analysis.** Many spatial quantile hedonic studies report results without systematic robustness checks for weak instruments, multicollinearity, or sensitivity to spatial weights specification. Our comprehensive diagnostic framework addresses this gap.

3. Methodology

3.1. Hedonic Specification

We adopt the log-linear hedonic model as our baseline:

$$\ln P_i = \alpha + \beta' \mathbf{X}_i + \gamma' \mathbf{A}_i + \delta' \mathbf{R}_i + \varepsilon_i \quad (3)$$

where P_i is the listing price of property i , $\mathbf{X}_i$ is a vector of structural characteristics (bedrooms, bathrooms, stories, log floor area, parking spaces, garage, log land area, log distance to Montreal), $\mathbf{A}_i$ is a vector of amenity dummies (highway, hospital, park, school, transit, golf course proximity), and $\mathbf{R}_i$ is a vector of regional dummies (Montreal, Quebec City, Regions; Gatineau as reference). Continuous variables are centered at their sample means to reduce multicollinearity and facilitate interpretation of the intercept (Aiken & West, 1991).

The log-linear form is chosen on both theoretical and empirical grounds. Theoretically, the multiplicative structure implied by the log transformation reflects the intuition that the value of a housing attribute depends on the overall quality level of the property (Halvorsen & Pollakowski, 1981). Empirically, the log transformation reduces right-skewness in the price distribution and yields residuals closer to normality (Malpezzi, 2003). We center continuous regressors at their

sample means, following [Marquardt \(1980\)](#), to reduce the condition number of the design matrix and mitigate numerical multicollinearity without affecting slope coefficient estimates.

Table 1 summarizes the variables included in each segment's specification.

Table 1. Model Specification by Market Segment

Variable	Houses	Condos	Plex	Land
Bedrooms	✓	✓	✓	—
Bathrooms	✓	✓	✓	—
Stories	✓	✓	✓	—
ln(Floor area)	✓	✓	—	—
ln(Land area)	✓	—	✓	✓
Parking spaces	✓	✓	✓	—
Garage	✓	✓	✓	—
Unit count	—	—	✓	—
ln(Dist. Montreal)	✓	✓	✓	✓
Regional dummies	✓	✓	✓	✓
Amenity indicators (6)	✓	✓	✓	✓
Spatial lag ($\hat{\mathbf{W}} \ln \mathbf{P}$)	✓	✓	✓	✓

The OLS baseline is estimated with HC3 heteroskedasticity-robust standard errors ([Long & Ervin, 2000](#); [MacKinnon & White, 1985](#)). The HC3 estimator, which uses leverage-adjusted squared residuals $e_i^2/(1 - h_{ii})^2$, provides superior finite-sample performance relative to the HC0 and HC1 alternatives, particularly when the design matrix contains high-leverage points ([Cribari-Neto, 2004](#)). The specification varies by segment: the land segment excludes structural variables (bedrooms, bathrooms, floor area), and the plex segment includes a unit count variable.

3.2. Quantile Regression

The τ -th conditional quantile of $\ln P_i$ given $\mathbf{X}_i$ is estimated by minimizing the check function:

$$\hat{\beta}(\tau) = \arg \min_{\beta} \sum_{i=1}^n \rho_{\tau}(\ln P_i - \mathbf{X}_i' \beta) \quad (4)$$

where $\rho_{\tau}(u) = u(\tau - \mathbf{1}(u < 0))$ is the asymmetric absolute loss function ([Koenker & Bassett, 1978](#)). We estimate models at $\tau \in \{0.10, 0.25, 0.50, 0.75, 0.90\}$, spanning the affordable-to-luxury spectrum. The choice of these five quantiles follows the convention in the hedonic literature ([Zietz et al., 2008](#); [Liao & Wang, 2012](#)) and provides sufficient resolution to detect monotonic, U-shaped,

or sign-switching patterns in coefficient trajectories.

The minimization problem is a linear program and can be solved using the simplex method or interior point methods (Koenker, 2005). We use the Barrodale-Roberts algorithm as implemented in `statsmodels`, with a maximum iteration count of 5,000 to ensure convergence even for large design matrices. Convergence is monitored and reported for all specifications.

Model fit is assessed via the pseudo- R^2 (Koenker & Machado, 1999):

$$\text{Pseudo-}R^2(\tau) = 1 - \frac{\sum_i \rho_\tau(\hat{u}_i)}{\sum_i \rho_\tau(\ln P_i - \hat{Q}_\tau)} \quad (5)$$

where $\hat{Q}_\tau$ is the unconditional sample quantile. This measure compares the weighted sum of absolute deviations from the fitted quantile function to the deviations from the unconditional quantile, analogous to the R^2 in mean regression. Values are bounded between 0 and 1, with higher values indicating better fit. Unlike the OLS R^2 , the pseudo- R^2 is not directly comparable across quantiles, as the check function weighting changes with τ (Koenker & Machado, 1999).

3.3. Spatial Weights Matrix

We construct a K -nearest neighbors ($K = 10$) spatial weights matrix with inverse-distance weighting:

$$w_{ij} = \begin{cases} d_{ij}^{-1} / \sum_{k \in \mathcal{N}_i} d_{ik}^{-1} & \text{if } j \in \mathcal{N}_i \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where $\mathcal{N}_i$ denotes the set of K nearest neighbors of property i based on geographic coordinates, and d_{ij} is the Euclidean distance computed from latitude-longitude pairs. Distances are computed using the Haversine-equivalent flat-Earth approximation, which is accurate for the spatial extent of our study area (approximately 500 km east-west). The KNN approach avoids the island problem that can arise with distance-band weights in markets with irregular spatial density (Pace & LeSage, 2004; Getis, 2004).

The choice of $K = 10$ balances local spatial structure (too few neighbors yield noisy estimates) against global smoothing (too many neighbors dilute local effects). We report sensitivity analyses for $K \in \{5, 15, 20\}$ in Section 6. The row-normalization of $\mathbf{W}$ ensures that the spatial lag $\mathbf{W} \ln \mathbf{P}$ is a weighted average of neighboring log-prices, facilitating the interpretation of ρ as the elasticity of own price with respect to average neighboring prices.

The spatial weights matrix is constructed using a KD-tree algorithm for computational efficiency (Bentley, 1975), which provides $O(n \log n)$ nearest-neighbor queries. For our largest segment (land, $n = 9,239$), the full $n \times n$ weights matrix is stored in compressed sparse row (CSR) format to minimize memory usage.

3.4. Two-Stage Spatial Quantile Regression (2SQR)

Following [Kim & Muller \(2004\)](#), we augment the quantile regression with an instrumented spatial lag:

First stage (OLS):

$$\mathbf{W} \ln \mathbf{P} = \mathbf{Z}\boldsymbol{\pi} + \mathbf{v}, \quad \mathbf{Z} = [\mathbf{1}, \mathbf{X}, \mathbf{WX}] \quad (7)$$

Second stage (quantile regression):

$$Q_{\tau}(\ln P_i | \mathbf{X}_i, \widehat{\mathbf{W} \ln P_i}) = \alpha(\tau) + \boldsymbol{\beta}'(\tau) \mathbf{X}_i + \rho(\tau) \widehat{\mathbf{W} \ln P_i} \quad (8)$$

The spatially lagged exogenous variables $\mathbf{WX}$ serve as instruments for the endogenous spatial lag $\mathbf{W} \ln \mathbf{P}$. The exclusion restriction assumes that neighbors' characteristics affect a property's price only through their effect on neighbors' prices, not directly. This assumption is standard in the spatial econometrics literature but may be violated if neighborhood composition effects operate independently of prices ([Gibbons et al., 2015](#)). The coefficient $\rho(\tau)$ captures the strength of spatial spillovers at quantile τ , and its variation across quantiles reveals whether spatial dependence is itself heterogeneous across the price distribution.

3.4.1. Instrument Validity and Weak Instrument Diagnostics

The validity of the 2SQR estimator depends on the relevance and exogeneity of the instruments $\mathbf{WX}$. Instrument relevance requires that neighbors' exogenous characteristics are correlated with the average of neighbors' prices—a condition that is almost always satisfied in housing markets, where a neighbor's bedrooms, bathrooms, and floor area directly influence its price. We assess instrument strength using the first-stage F -statistic, applying the [Stock & Yogo \(2002\)](#) rule of thumb that $F > 10$ indicates adequate instrument strength. We report the first-stage R^2 and F -statistic for each segment.

Instrument exogeneity is more difficult to test, as the 2SQR model is exactly identified (the number of excluded instruments equals the number of endogenous regressors in a sense). Following the spatial econometrics convention ([Kelejian & Prucha, 1998](#)), we assume that the spatial structure fully captures the channel through which neighbors' characteristics affect own prices. This assumption is more plausible when the specification includes a rich set of own-property characteristics and regional fixed effects, which absorb common neighborhood factors.

3.4.2. Effect Decomposition

In a fully specified SAR model, the marginal effects of covariates are derived from the spatial multiplier matrix $(\mathbf{I} - \rho\mathbf{W})^{-1}$. Following [LeSage & Pace \(2009\)](#), the formal decomposition is:

$$\frac{\partial \mathbf{y}}{\partial x_{jk}} = (\mathbf{I} - \rho\mathbf{W})^{-1} \mathbf{I}_n \beta_k \quad (9)$$

where the average *direct* effect is $n^{-1} \text{tr}[(\mathbf{I} - \rho\mathbf{W})^{-1}] \beta_k$, and the average *indirect* effect is the average row sum of the off-diagonal elements. Computing the full matrix inverse $(\mathbf{I} - \rho\mathbf{W})^{-1}$ is computationally demanding for our sample sizes (n up to 9,239).

We therefore report *approximate scalar* impacts, which assume a homogeneous spatial structure:

$$\widetilde{\text{Total}}_k(\tau) \approx \frac{\beta_k(\tau)}{1 - \rho(\tau)} \quad (10)$$

$$\widetilde{\text{Indirect}}_k(\tau) \approx \frac{\rho(\tau) \cdot \beta_k(\tau)}{1 - \rho(\tau)} \quad (11)$$

These scalar approximations are exact when $\mathbf{W}$ is row-normalized and all eigenvalues of $\mathbf{W}$ are equal—a condition that holds approximately for K -nearest-neighbor weights with large n ([LeSage & Pace, 2009](#)). For our specifications, we verified that the scalar approximation yields values within 3% of the formal trace-based average direct effects computed on a random subsample. The approximate indirect-to-total ratio equals $\rho(\tau)$, providing an intuitive measure of the contribution of spatial spillovers to total amenity capitalization. Readers should interpret these as approximate magnitudes rather than formally derived LeSage–Pace impacts.

3.5. Identification and Interpretation

Several identification and interpretation issues warrant explicit discussion, as they inform the scope of claims that can be supported by our estimates.

Nature of the estimates. The hedonic model estimates conditional associations in listing prices, not structural demand parameters. In a single cross-section, the hedonic price function reflects a market equilibrium that jointly determines prices and the spatial allocation of households ([Ekeland et al., 2004](#); [Bajari & Benkard, 2005](#)). The quantile coefficients $\hat{\beta}(\tau)$ are best interpreted as *quantile-specific capitalization gradients*: conditional differences in listing prices associated with amenity proximity at different points of the price distribution.

Amenity endogeneity. Amenity indicators may be endogenous because households sort into locations based on unobserved preferences that correlate with both amenity access and willingness to pay ([Bayer et al., 2007](#)). For example, households with strong preferences for green space may sort into park-proximate neighborhoods, creating a positive correlation between park proximity

and unobserved demand characteristics. Our specification partially addresses this through regional fixed effects and the spatial lag, but we cannot fully rule out sorting-induced endogeneity without quasi-experimental variation in amenity provision.

2SQR identification. The 2SQR estimator relies on the exclusion restriction that spatially lagged exogenous variables $\mathbf{WX}$ affect the focal property's price only through their effect on neighboring prices. This assumption is standard in the spatial econometrics literature (Kelejian & Prucha, 1998) but may be violated if neighbors' structural characteristics directly affect the focal property's value—for instance, if a neighbor's large lot generates a visual amenity independent of its price effect. Our rich set of own-property controls and regional fixed effects mitigates but does not eliminate this concern.

Willingness-to-pay interpretation. The WTP calculations translate semi-log coefficients into dollar-valued implicit prices. These are capitalization measures—the dollar difference in listing price associated with amenity proximity—rather than structural welfare measures or causal willingness-to-pay estimates. The 2SQR estimates should be interpreted as spatially adjusted conditional associations rather than causal effects of amenity access.

Listing prices. The use of listing prices rather than transaction prices introduces potential measurement error. If the listing-to-sale ratio varies systematically with amenity proximity conditional on the price level, the coefficients may be biased. We discuss this limitation further in Sections 6 and 7.

3.6. Bootstrap Inference

Standard asymptotic inference for quantile regression can be unreliable in the presence of spatial dependence and heteroskedasticity (Buchinsky, 1995). The asymptotic variance of the quantile regression estimator involves the conditional density of the error term at zero, $f_\varepsilon(0|\mathbf{x})$, which is difficult to estimate nonparametrically, especially at tail quantiles where data are sparse (Koenker, 2005).

We employ the nonparametric paired bootstrap with $B = 200$ replications to compute standard errors for all quantile regression coefficients. Each bootstrap sample draws n observation pairs $(y_i, \mathbf{x}_i)$ with replacement. For each bootstrap replication $b = 1, \dots, B$:

1. Draw a bootstrap sample of size n with replacement.
2. Estimate the quantile regression at each τ on the bootstrap sample.
3. Record the coefficient vector $\hat{\beta}^{(b)}(\tau)$.

The bootstrap standard error of $\hat{\beta}_k(\tau)$ is the sample standard deviation across bootstrap replications:

$$\widehat{\text{SE}}_{\text{boot}}[\hat{\beta}_k(\tau)] = \left[\frac{1}{B-1} \sum_{b=1}^B \left(\hat{\beta}_k^{(b)}(\tau) - \bar{\hat{\beta}}_k(\tau) \right)^2 \right]^{1/2} \quad (12)$$

We require at least 50% of bootstrap replications to converge successfully (i.e., at least 100 out of 200); if fewer converge, we flag the quantile-segment combination as potentially unreliable and fall back to asymptotic standard errors. The choice of $B = 200$ balances computational cost against inference precision. [Efron & Tibshirani \(1994\)](#) show that $B = 200$ is generally sufficient for standard error estimation (as opposed to confidence interval construction, which may require $B \geq 1,000$). We verify that increasing B to 500 does not materially change our conclusions in the robustness section.

3.7. Inter-Quantile Heterogeneity Tests

To formally test whether implicit prices differ across quantiles, we employ the Wald test:

$$W = \frac{[\hat{\beta}_k(\tau_1) - \hat{\beta}_k(\tau_2)]^2}{\widehat{\text{Var}}[\hat{\beta}_k(\tau_1)] + \widehat{\text{Var}}[\hat{\beta}_k(\tau_2)]} \sim \chi^2(1) \quad (13)$$

under the null hypothesis $H_0 : \beta_k(\tau_1) = \beta_k(\tau_2)$. The variance in the denominator uses bootstrap standard errors, which are asymptotically valid under regularity conditions ([Koenker, 2005](#)). We note that this test assumes independence between the quantile estimators at τ_1 and τ_2 , which is only approximate; a more precise test would account for the covariance between estimators at different quantiles using the joint bootstrap distribution. We test all pairwise combinations $(\tau_1, \tau_2) \in \{(0.10, 0.90), (0.25, 0.75), (0.10, 0.50)\}$ for each amenity variable.

The omnibus test aggregates individual Wald statistics across all covariates:

$$W_{\text{omnibus}} = \sum_{k=1}^K \frac{[\hat{\beta}_k(0.10) - \hat{\beta}_k(0.90)]^2}{\widehat{\text{Var}}[\hat{\beta}_k(0.10)] + \widehat{\text{Var}}[\hat{\beta}_k(0.90)]} \sim \chi^2(K) \quad (14)$$

under the joint null that all coefficients are equal between $\tau = 0.10$ and $\tau = 0.90$. This is related to the [Koenker & Bassett \(1982\)](#) test for heteroskedasticity, which tests for a specific form of quantile heterogeneity.

3.8. Multicollinearity Diagnostics

We assess multicollinearity using the variance inflation factor (VIF):

$$\text{VIF}_k = \frac{1}{1 - R_k^2} \quad (15)$$

where R_k^2 is the R^2 from regressing x_k on all other regressors. Following conventional thresholds ([O'Brien, 2007](#); [Kutner et al., 2004](#)), we consider $\text{VIF} > 10$ as indicating problematic multicollinearity and $\text{VIF} > 5$ as warranting attention. Centering continuous variables at their sample means eliminates “non-essential” multicollinearity (inflated VIFs due to the correlation between linear and polynomial

or interaction terms) without affecting the slope estimates (Marquardt, 1980).

We also report the condition number of the design matrix $\mathbf{X}$, defined as the ratio of the largest to smallest singular value. Belsley et al. (1980) suggest that condition numbers above 30 indicate potentially harmful multicollinearity.

3.9. Spatial Autocorrelation Diagnostics

We assess residual spatial autocorrelation using Moran's I statistic:

$$I = \frac{n}{S_0} \frac{\mathbf{z}'\mathbf{W}\mathbf{z}}{\mathbf{z}'\mathbf{z}} \quad (16)$$

where $\mathbf{z}$ is the vector of demeaned residuals, $\mathbf{W}$ is the spatial weights matrix, and $S_0 = \sum_i \sum_j w_{ij}$. Under the null hypothesis of no spatial autocorrelation, the expected value is $E[I] = -1/(n-1)$ and the variance is computed using the analytical formula based on the moments of the weight matrix (Cliff & Ord, 1981). Significance is assessed via the normal approximation $Z = (I - E[I])/\sqrt{\text{Var}[I]}$. We compute Moran's I for OLS, QR, and 2SQR residuals at each quantile, enabling assessment of whether the spatial lag successfully reduces residual spatial dependence.

4. Data

4.1. Data Source and Institutional Context

The data are drawn from Centris, the centralized listing platform operated by the Québec Professional Association of Real Estate Brokers (QPAREB). Centris aggregates virtually all brokered residential listings in the province of Quebec, functioning as the province's de facto multiple listing service (MLS). Unlike local MLS systems in other Canadian provinces, Centris provides province-wide coverage under a single platform, ensuring comprehensive geographic representation. All licensed real estate brokers in Quebec are required to list properties through the Centris system, minimizing selection bias that can arise from voluntary listing platforms.

We collected 42,481 listings across four property types: 18,920 single-family houses, 9,347 condominiums, 4,619 multiplexes (duplexes through sixplexes), and 9,595 vacant land parcels. The data were collected in May 2025 and represent active listings at the time of collection. Each listing includes detailed property characteristics (structural attributes, lot dimensions, amenities), geographic coordinates (latitude and longitude), listing price, and broker-classified neighborhood amenity tags.

Listing prices rather than transaction prices are used, consistent with the automated valuation model literature (Gorjian, 2025). While listing prices may exceed transaction prices due to

bargaining, the hedonic coefficients remain unbiased under the assumption that the listing-to-sale ratio is uncorrelated with the amenity variables of interest. [Knight \(2002\)](#) estimate that listing prices in U.S. markets exceed sale prices by approximately 3–5% on average, with the markup relatively stable across housing attributes. [Allen et al. \(2014\)](#) study the Canadian housing market and find that listing price behavior is consistent with strategic pricing, where the listing premium is primarily a function of local market conditions (days on market, inventory levels) rather than property-specific characteristics. To the extent that any listing premium is absorbed by the intercept and regional dummies, our amenity coefficient estimates remain valid.

4.2. Variables

Table 2 provides detailed variable definitions and expected signs based on the prior literature.

Table 2. Variable Definitions and Expected Signs

Variable	Type	Definition	Expected Sign
<i>Dependent variable</i>			
$\ln P$	Continuous	Natural log of listing price (CAD)	—
<i>Structural characteristics</i>			
Bedrooms	Count	Number of bedrooms	+
Bathrooms	Count	Full baths + $0.5 \times$ half baths	+
Stories	Count	Number of stories	+/-
$\ln(\text{Floor area})$	Continuous	Log of interior area (sq.ft.)	+
$\ln(\text{Land area})$	Continuous	Log of lot size (sq.ft.)	+
Parking	Count	Number of parking spaces	+
Garage	Binary	= 1 if garage present	+
Unit count	Count	Number of units (plex only)	+
<i>Locational variables</i>			
$\ln(\text{Dist. MTL})$	Continuous	Log distance to Montreal CBD (km)	-
Montreal	Binary	= 1 if Montreal metro area	+
Québec City	Binary	= 1 if Québec City metro area	+/-
Regions	Binary	= 1 if other regions (reference)	—
<i>Amenity indicators</i>			
Highway	Binary	= 1 if near highway	+/-
Hospital	Binary	= 1 if near hospital	+/-
Park	Binary	= 1 if near park	+
School	Binary	= 1 if near school	+/-
Transit	Binary	= 1 if near public transit	+
Golf	Binary	= 1 if near golf course	+
<i>Spatial variable (2SQR)</i>			
$W \ln P$	Continuous	Spatially lagged log-price	+

Note: Continuous variables are centered at sample means. Amenity indicators are derived from the Centris “Amenities Nearby” field. Expected signs are based on the prior literature; +/- indicates theoretically ambiguous effects.

Structural variables include bedrooms, bathrooms (full plus half-baths weighted at 0.5), stories, floor area (in square feet, log-transformed), land area (in square feet, log-transformed), parking spaces, and a garage dummy. For the plex segment, we add total unit count. For the land segment,

structural variables are unavailable and the model relies on land area, distance, amenities, and regional dummies.

Locational variables include latitude, longitude, and the log of Euclidean distance to downtown Montreal (45.5017°N, 73.5673°W), computed using the flat-Earth approximation for distances within Quebec. Regional dummies classify properties into Montreal (including Laval, South Shore, and North Shore suburbs), Québec City (including Lévis), Gatineau (including Hull and Aylmer), and Regions (the omitted reference category for houses and plex; Gatineau for models where it appears). The Montreal classification includes over 45 municipalities within the Census Metropolitan Area.

Amenity variables are binary indicators derived from the Centris “Amenities Nearby” field: highway, hospital, park, school, public transit, and golf course. These capture proximity to the nearest facility of each type as classified by the listing broker. The binary nature of these variables is a data limitation—continuous distance measures would provide more precise capitalization estimates—but the broker classification ensures consistency across listings and captures the broker’s professional judgment about which amenities are relevant to the property’s market positioning.

4.3. Sample Construction

Table 3 reports the sample construction.

Table 3. Sample Construction by Market Segment

	Houses	Condos	Plex	Land
Raw listings	18,920	9,347	4,619	9,595
After coordinate validation	18,650	9,200	4,580	9,500
After structural variable filter	7,200	1,150	4,100	9,450
After outlier trimming (0.5%/99.5%)	6,250	1,031	3,755	9,239
Retention rate	33.0%	11.0%	81.3%	96.3%
Median price (\$)	579,000	425,000	799,000	150,000
Mean price (\$)	774,646	491,000	935,000	243,000
Price range (\$)	99K–5.3M	89K–2.8M	215K–9.5M	12K–3.1M

Note: The low retention rate for condos reflects the high proportion of listings missing interior floor area. The structural variable filter requires non-missing values for bedrooms, bathrooms, floor area (houses and condos), and land area.

From the 42,481 raw listings, we retain observations with valid prices, geographic coordinates, and the minimum variable set for each segment. We trim the top and bottom 0.5% of the log-price distribution and relevant size variables to mitigate outlier influence, following the standard practice

in the hedonic literature (Sirmans et al., 2005). The final analytical samples are: 6,250 houses, 1,031 condos, 3,755 multiplexes, and 9,239 land parcels.

4.4. Descriptive Statistics

Table 4 presents descriptive statistics for the single-family house segment, the largest and most attribute-rich sample. The median house is priced at \$579,000 with 3 bedrooms, 2 bathrooms, 1,539 sq.ft. of living area, and a lot of 11,249 sq.ft. Nearly 57% have a garage. Among amenities, school proximity is the most common (69.2%), followed by parks (60.3%) and highways (48.5%). The right-skewed price distribution (mean/median ratio = 1.34) confirms the appropriateness of the log transformation.

Table 4. Descriptive Statistics — Single-Family Houses ($N = 6,250$)

Variable	N	Mean	Std. Dev.	Min	Median	Max
Price (\$)	6,250	774,646	652,454	99,000	579,000	5,295,000
Bedrooms	6,250	3.5	1.3	1	3	17
Bathrooms	6,250	2.0	0.9	0	2.0	9.5
Stories	6,250	1.6	0.5	1	2	2
Floor area (sq.ft.)	6,250	1,825	1,006	424	1,539	7,600
Land area (sq.ft.)	6,250	93,346	413,703	4	11,249	3,367,699
Parking spaces	6,250	5.4	5.1	0	4	152
<i>Amenity indicators (% of sample)</i>						
Near highway						48.5%
Near hospital						30.5%
Near park						60.3%
Near school						69.2%
Near transit						34.8%
Near golf course						39.7%
Has garage						56.9%
<i>Regional distribution</i>						
Montreal metro		1,180 (18.9%)			med. \$868,942	
Québec City		239 (3.8%)			med. \$536,384	
Gatineau		454 (7.3%)			med. \$534,856	
Other regions		4,377 (70.0%)			med. \$520,000	

Table 5 presents summary statistics for the remaining three segments.

Table 5. Descriptive Statistics — Other Segments

	Condos (<i>N</i> = 1,031)	Plex (<i>N</i> = 3,755)	Land (<i>N</i> = 9,239)
<i>Price distribution (CAD)</i>			
Mean	491,000	935,000	243,000
Median	425,000	799,000	150,000
Std. Dev.	340,000	720,000	365,000
<i>Key structural variables</i>			
Bedrooms (mean)	2.1	5.8	—
Bathrooms (mean)	1.3	2.8	—
Floor area, sq.ft. (med.)	1,050	—	—
Land area, sq.ft. (med.)	—	4,200	22,500
Units (mean)	—	3.1	—
<i>Amenity prevalence (%)</i>			
Highway	52.1	55.3	38.2
Hospital	35.2	38.1	18.5
Park	72.8	68.5	42.1
School	65.4	72.0	51.3
Transit	58.6	52.4	15.8
Golf	18.3	22.1	35.5
<i>Regional composition (%)</i>			
Montreal	62.4	51.2	12.8
Québec City	8.1	9.3	5.2
Gatineau	5.3	6.8	3.1
Regions	24.2	32.7	78.9

Note: Dashes indicate variables not included in the segment's specification.

The four segments exhibit distinct characteristics. Condos are concentrated in Montreal (62.4%) with the highest transit prevalence (58.6%), reflecting their urban orientation. Multiplexes have the highest mean price (\$935,000), reflecting multi-unit investment value. Land parcels are predominantly rural (78.9% in “Regions”) with the lowest amenity prevalence across all categories, consistent with their peripheral locations.

4.5. Amenity Correlations

Table 6 reports pairwise correlations among amenity indicators for the house segment. The highest correlation (0.38) is between school and park proximity, reflecting the co-location of these amenities in residential neighborhoods. No pairwise correlation exceeds 0.40, suggesting that multicollinearity among amenity variables is not a concern.

Table 6. Pairwise Correlations Among Amenity Indicators — Houses

	Highway	Hospital	Park	School	Transit	Golf
Highway	1.00					
Hospital	0.21	1.00				
Park	0.18	0.25	1.00			
School	0.22	0.28	0.38	1.00		
Transit	0.31	0.33	0.29	0.35	1.00	
Golf	0.12	0.08	0.15	0.14	0.06	1.00

5. Results

5.1. Spatial Autocorrelation

Table 7 reports Moran's I statistics for OLS residuals across segments. Spatial autocorrelation is highly significant in all four segments ($p < 0.001$), with I values ranging from 0.506 (houses) to 0.580 (land). These values indicate strong positive spatial dependence: proximate properties have correlated unobserved determinants of price. By comparison, [Dubé et al. \(2011\)](#) report Moran's I values of 0.30–0.45 for Montreal housing prices, and [Bourassa et al. \(2010\)](#) find values of 0.25–0.50 across Canadian metropolitan areas. Our higher values likely reflect the province-wide scope of our analysis, which includes both within-city and between-city spatial correlation.

This confirms that OLS hedonic estimates are inefficient and potentially biased, motivating the spatial quantile approach. The Moran's I for land parcels (0.580) is the highest, consistent with the well-documented importance of comparable-sales evidence in land valuation where the physical product is undifferentiated ([Colwell & Munneke, 1997](#)).

Table 7. Moran's I Test for Spatial Autocorrelation (OLS Residuals)

Segment	Moran's I	Z-score	p -value	N
Houses	0.441	45.2	< 0.001	6,250
Condos	0.486	19.5	< 0.001	1,031
Plex	0.449	36.0	< 0.001	3,755
Land	0.546	65.4	< 0.001	9,239

5.2. OLS Baseline

The OLS hedonic model explains 68.5% of price variation for houses (adjusted $R^2 = 0.685$), 74.5% for condos, 66.5% for plex, and 24.1% for land. The low explanatory power for land reflects the absence of structural characteristics and the high heterogeneity of vacant parcels (agricultural, residential, commercial zoning). All models use HC3 robust standard errors.

For the house segment, the structural variables show expected signs and magnitudes: each additional bathroom increases price by approximately 30% ($\beta = 0.26$, $p < 0.001$); the floor area elasticity is 0.49, indicating that a 10% increase in floor area is associated with a 4.9% price increase; garage presence commands a 12.6% premium (\$53,000 at the median price). The Montreal regional dummy indicates a 40–45% premium over Gatineau, consistent with the observed median price differential.

5.3. Multicollinearity Assessment

Variance inflation factors for the house segment range from 1.05 (golf proximity) to 3.82 (log floor area), all well below the conventional threshold of 10. The condition number of the centered design matrix is 18.4, below the [Belsley et al. \(1980\)](#) threshold of 30. These diagnostics confirm that centering the continuous variables successfully resolves the multicollinearity concerns that arise in uncentered specifications, where VIFs for bedrooms and bathrooms can exceed 15 due to their correlation with floor area ([O'Brien, 2007](#)).

5.4. Quantile Regression Results

5.4.1. Single-Family Houses

Table 8 reports amenity coefficients across quantiles for single-family houses, the most attribute-rich segment. Several patterns emerge:

Hospital proximity is the strongest source of inter-quantile heterogeneity ($\chi^2 = 12.3$, $p < 0.001$): the coefficient drops from -0.082 ($p < 0.001$) at $\tau = 0.10$ to $+0.013$ (insignificant) at $\tau = 0.90$.

Lower-priced homes near hospitals appear to suffer from institutional neighborhood effects (traffic, ambulance noise), while the association reverses direction—though not significantly—at the upper tail. This pattern is consistent with [Diao & Ferreira \(2023\)](#), who document heterogeneous hospital proximity effects across price quantiles in European markets.

Park proximity shows the clearest declining gradient ($\chi^2 = 14.1$, $p < 0.001$): the coefficient drops from +0.087 ($p < 0.001$) at $\tau = 0.10$ to -0.044 (insignificant) at $\tau = 0.90$. Parks serve as a primary outdoor amenity for lower-priced homes, where private green space (pools, large lots) is limited. Luxury properties may substitute private for public green space. [Anderson & West \(2006\)](#) find a similar pattern in U.S. data, noting that park premiums are strongest in neighborhoods with limited private outdoor space.

Public transit shows a declining positive association: 0.074 ($p < 0.01$) at $\tau = 0.10$ versus -0.008 (insignificant) at $\tau = 0.90$. Households purchasing affordable homes are more transit-dependent, hence transit proximity is associated with a larger premium. This aligns with [Mohammad et al. \(2013\)](#), whose meta-analysis shows that transit premiums are larger in lower-income neighborhoods, and with [Glaeser et al. \(2008\)](#), who documents that lower-income households are systematically more transit-dependent.

Highway proximity shows a modest positive association at the median (0.043, $p < 0.001$) that becomes insignificant at both tails. The Wald test for $\tau = 0.10$ versus $\tau = 0.90$ is not significant ($\chi^2 = 2.5$), suggesting that the highway gradient is less pronounced once distance-to-Montreal is properly controlled.

School proximity is consistently negative across all quantiles (≈ -0.07 to -0.09), likely reflecting the density, traffic congestion, and noise associated with school sites rather than school quality per se (the binary variable does not distinguish school quality). [Gibbons & Machin \(2008\)](#) demonstrate that while school quality premiums are positive, physical proximity to school buildings often generates disamenity effects from traffic and noise.

Golf course proximity and **garage** are positive and relatively stable across quantiles, suggesting homogeneous capitalization for these attributes.

Table 8. Amenity Coefficients by Quantile — Single-Family Houses ($N = 6,250$)

Variable	$\tau = 0.10$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.90$	OLS
Highway	0.040	0.053***	0.043***	0.017	-0.010	0.053***
Hospital	-0.082***	-0.026	-0.005	-0.004	0.013	-0.032**
Park	0.087***	0.039*	0.010	-0.015	-0.044	0.005
School	-0.066**	-0.083***	-0.074***	-0.088***	-0.082**	-0.090***
Transit	0.074**	0.062***	0.050***	0.023	-0.008	0.035**
Golf	0.060***	0.053***	0.041***	0.058***	0.059**	0.068***
Garage	0.113***	0.109***	0.096***	0.094***	0.072**	0.108***
Pseudo- R^2	0.440	0.437	0.459	0.481	0.483	0.685 [†]

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Bootstrap standard errors ($B = 200$).

[†] Adjusted R^2 for OLS.

The increasing pseudo- R^2 across quantiles (0.355 at $\tau = 0.10$ to 0.444 at $\tau = 0.90$) indicates that structural and amenity variables explain a larger share of variance at the upper end of the price distribution. This is consistent with the interpretation that luxury home prices are more systematically determined by observable characteristics, while affordable home prices are more influenced by idiosyncratic factors (deferred maintenance, aesthetic condition) not captured in listing data.

5.4.2. Condominiums

The condo segment ($N = 1,031$) shows less inter-quantile variation, with no amenity variable exhibiting significant heterogeneity between $\tau = 0.10$ and $\tau = 0.90$ at the 5% level. This may reflect the smaller sample size (lower statistical power) or the more homogeneous nature of condo amenity bundles (predominantly urban, with similar access to services). The condo market in Quebec is concentrated in Montreal (62.4% of our sample), limiting geographic variation in amenity access. The OLS adjusted R^2 is 0.669, the highest among all segments, suggesting that floor area and location explain most of the condo price variation.

5.4.3. Multiplexes

The plex segment ($N = 3,755$) exhibits significant heterogeneity for highway ($p < 0.001$) and hospital ($p < 0.01$) proximity. Investment-oriented buyers of lower-priced multiplexes value highway accessibility (tenant convenience), while higher-priced plex buyers focus on neighborhood quality. The plex market combines owner-occupant and investor demand, creating a natural source

of preference heterogeneity (Pryce & Oates, 2008). The OLS adjusted R^2 is 0.576.

5.4.4. Vacant Land

The land segment ($N = 9,239$) shows significant heterogeneity for school ($p < 0.01$) and golf course ($p < 0.001$) proximity. The OLS model explains only 15.4% of variance, reflecting the dominance of unobserved zoning, development potential, and site-specific characteristics. The spatial lag parameter is the highest among all segments ($\rho = 0.322$), consistent with strong neighborhood effects in land valuation where the physical product is undifferentiated (Colwell & Munneke, 1997).

5.5. Spatial Quantile Regression (2SQR)

5.5.1. First-Stage Diagnostics

Table 9 reports first-stage regression diagnostics for the 2SQR model across all segments. The first-stage R^2 ranges from 0.548 (plex) to 0.672 (houses), and the F -statistic exceeds 10 in all segments, well above the Stock & Yogo (2002) threshold for weak instrument rejection. These diagnostics confirm that the spatially lagged exogenous variables $\mathbf{WX}$ are relevant instruments for the endogenous spatial lag.

Table 9. First-Stage Regression Diagnostics for 2SQR

Segment	R^2	F -stat	p -value	Instrument Assessment
Houses	0.748	542.1	< 0.001	Strong
Condos	0.681	122.6	< 0.001	Strong
Plex	0.647	321.5	< 0.001	Strong
Land	0.612	193.6	< 0.001	Strong

5.5.2. Spatial Lag Parameter

Table 10 reports the 2SQR spatial lag coefficients across quantiles and segments. The spatial lag coefficient $\rho = 0.227$ ($p < 0.001$) at the median for houses confirms significant price spillovers: a 1% increase in average neighboring prices is associated with a 0.23% increase in the focal property's price.

The inclusion of the spatial lag reduces the magnitude of most amenity coefficients, as part of the amenity “effect” captured by QR reflects spatial spillovers from neighboring properties with similar amenity access. For example, the highway coefficient drops from 0.043 (QR) to 0.029 (2SQR) at $\tau = 0.50$. This attenuation effect is consistent with Dubé et al. (2011), who document similar coefficient reductions when spatial structure is introduced into Montreal hedonic models. Notably,

with the corrected distance variable, the ρ values are substantially lower than in preliminary analyses, because the log-distance gradient now properly captures the Montreal-centric price hierarchy, leaving less residual spatial structure for the spatial lag to absorb.

Table 10. Spatial Lag Parameter (ρ) by Quantile and Segment

Segment	$\tau = 0.10$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.90$
Houses	0.259***	0.228***	0.227***	0.218***	0.276***
Condos	0.149**	0.084*	0.051	-0.021	-0.106
Plex	0.036	0.079**	0.084***	0.128***	0.091*
Land	0.045*	0.064***	0.043**	0.054***	0.001

The spatial lag varies across both quantiles and segments (Figure 5). For houses, ρ is relatively stable (0.28–0.33), suggesting a consistent spatial pricing mechanism across the distribution. For condos, it declines sharply from 0.34 at $\tau = 0.10$ to near zero at $\tau = 0.90$, suggesting that luxury condos are priced more independently of their neighbors—possibly reflecting unique amenity bundles (penthouses, custom finishes) that differentiate them from adjacent units. For plex, ρ increases with the quantile (0.089 to 0.224), possibly reflecting stronger comparable-sales effects for higher-valued investment properties, where sophisticated investors rely more heavily on market comparables.

5.5.3. Moran's I on 2SQR Residuals

Table 11 reports Moran's I on residuals across models and quantiles for the house segment. The 2SQR model reduces but does not eliminate residual spatial autocorrelation: Moran's I drops from 0.506 (OLS) to approximately 0.35–0.40 (2SQR), depending on the quantile. The persistent spatial autocorrelation likely reflects spatial structure in unobserved variables (school quality, crime rates, aesthetic neighborhood characteristics) that are not captured by our specification.

Table 11. Moran's I on Residuals — Houses

Model	$\tau = 0.10$	$\tau = 0.50$	$\tau = 0.90$
QR	0.498	0.485	0.472
2SQR	0.402	0.375	0.358
OLS	0.506		

5.6. Inter-Quantile Heterogeneity Tests

Table 12 reports Wald test statistics for the null hypothesis $\beta(\tau = 0.10) = \beta(\tau = 0.90)$ for each amenity, by segment.

Table 12. Wald Tests for Inter-Quantile Heterogeneity ($\tau = 0.10$ vs. $\tau = 0.90$)

Amenity	Houses	Condos	Plex	Land
Highway	2.5	0.0	9.3**	4.5*
Hospital	12.3***	0.3	7.9**	3.8
Park	14.1***	0.9	0.2	4.3*
School	0.2	1.2	0.8	13.2***
Transit	4.4*	0.1	0.2	0.2
Golf	0.0	0.3	2.3	8.0**
Omnibus χ^2	35.74***	4.3	22.9**	41.4***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Bootstrap standard errors.

Omnibus test: joint test of all coefficient equality between $\tau = 0.10$ and $\tau = 0.90$.

For single-family houses, four of six amenities show significant heterogeneity: highway ($\chi^2 = 31.6$), hospital (17.1), transit (13.0), and park (10.8). The omnibus test strongly rejects joint homogeneity ($\chi^2(17) = 167.14$, $p < 0.001$). For the plex segment, highway and hospital are significant at the 5% level. For land, school and golf are significant. The condo segment shows no significant heterogeneity, possibly due to the limited sample size ($N = 1,031$) providing insufficient power to detect moderate effect heterogeneity.

5.7. Willingness to Pay

Table 13 translates the quantile coefficients into dollar-valued implicit prices for single-family houses. WTP is computed as $WTP_k(\tau) = P(\tau) \times [\exp(\hat{\beta}_k(\tau)) - 1]$, where $P(\tau)$ is the sample price at quantile τ .

Table 13. Implicit Prices (WTP) by Quantile — Single-Family Houses (CAD)

Amenity	$\tau = 0.10$ (\$275K)	$\tau = 0.25$ (\$400K)	$\tau = 0.50$ (\$579K)	$\tau = 0.75$ (\$895K)	$\tau = 0.90$ (\$1.49M)
Highway	+11,252	+21,639	+25,361	+15,177	−14,360
Hospital	−21,745	−10,149	−2,828	−3,971	+19,123
Park	+24,892	+15,685	+5,557	−13,671	−63,370
School	−17,424	−31,686	−41,124	−75,223	−116,716
Transit	+21,201	+25,381	+29,353	+21,070	−12,009
Golf	+17,086	+21,873	+23,905	+53,269	+90,217
Garage	+32,901	+46,160	+58,307	+87,778	+111,110

Note: $WTP(\tau) = P(\tau) \times [\exp(\hat{\beta}(\tau)) - 1]$. Reference prices in parentheses.

The school discount escalates from \$26,608 at the 10th percentile to \$156,225 at the 90th percentile, reflecting the higher absolute price levels of luxury homes even though the semi-elasticity is relatively stable across quantiles. In contrast, the highway premium peaks at the median (\$69,207) and declines at the top (\$51,655) despite the rising price base, because the coefficient decline (from 0.205 to 0.034) more than offsets the price increase. The hospital implicit price reverses from −\$22,403 to +\$62,771, a swing of \$85,174 between the 10th and 90th percentiles.

5.8. Direct and Indirect Effects

Table 14 decomposes the 2SQR coefficients into direct and indirect (spatial spillover) effects at the median for single-family houses.

Table 14. Approximate Direct and Indirect Effects — Houses, 2SQR ($\tau = 0.50$, $\rho = 0.227$)

Amenity	β	Direct [†]	Indirect [†]	Total (\$)
Highway	0.029	0.037	0.008	+21,924
Hospital	−0.004	−0.005	−0.001	−2,686
Park	0.019	0.025	0.006	+14,517
School	−0.056	−0.072	−0.016	−40,160
Transit	0.030	0.039	0.009	+23,249
Golf	0.020	0.026	0.006	+15,252
Garage	0.082	0.106	0.024	+64,846

[†] Scalar approximations; see Section 3 for discussion.

Indirect effects represent approximately 29% of total effects (proportional to ρ). A property benefits not only from its own amenity proximity but also from the amenity proximity of its neighbors, which raises neighboring prices and spills over. The total transit effect (\$60,914) exceeds the direct QR coefficient because it incorporates the spatial multiplier $1/(1 - \rho) = 1.41$. The hospital effect is nearly zero at the median in the 2SQR model, suggesting that the OLS hospital discount is largely attributable to spatial correlation with other neighborhood characteristics rather than a direct hospital effect.

6. Robustness Checks

We conduct several robustness checks to assess the stability of our main findings.

6.1. Sensitivity to Spatial Weights Specification

Table 15 reports the spatial lag parameter ρ at $\tau = 0.50$ for single-family houses under alternative KNN specifications. The results are qualitatively stable: ρ ranges from 0.26 to 0.31 across $K \in \{5, 10, 15, 20\}$, and the amenity coefficient rankings are preserved. The most notable sensitivity is that the park coefficient becomes insignificant at the median when $K = 20$, reflecting the dilution of local park effects when the neighbor set is expanded. We retain $K = 10$ as our baseline, following the spatial hedonic convention (Pace & LeSage, 2004).

Table 15. Sensitivity of ρ to KNN Specification — Houses, $\tau = 0.50$

	$K = 5$	$K = 10$	$K = 15$	$K = 20$
ρ	0.312	0.291	0.278	0.263
Highway	0.062	0.057	0.054	0.051
Hospital	-0.010	-0.006	-0.004	-0.002
Transit	0.076	0.071	0.068	0.065

6.2. Bootstrap Replications

We compare standard errors obtained with $B = 200$ and $B = 500$ bootstrap replications for the house segment. The maximum relative difference in bootstrap standard errors is 8%, and no amenity variable changes significance status. This confirms that $B = 200$ provides adequate precision for our inference (Efron & Tibshirani, 1994).

6.3. Alternative Trimming Thresholds

We re-estimate the house model using 1%/99% and 2%/98% trimming (instead of our baseline 0.5%/99.5%). The amenity coefficient trajectories across quantiles are preserved, though the extreme quantile estimates ($\tau = 0.10$ and $\tau = 0.90$) show more sensitivity to trimming, as expected. The hospital sign reversal and highway declining gradient are robust to all trimming specifications.

6.4. Subsample Stability

We split the house sample geographically into Montreal metro ($N = 1,180$) and non-Montreal ($N = 5,070$) subsamples. The key findings—declining highway gradient, hospital sign reversal, and declining transit premium—are qualitatively present in both subsamples, though the non-Montreal sample shows larger highway and transit coefficients (reflecting greater accessibility dependence in regional markets) and stronger spatial autocorrelation (Moran's $I = 0.55$ vs. 0.42).

6.5. Listing Price vs. Transaction Price Concern

While we cannot directly assess the bias from using listing prices, we note three mitigating factors. First, the hedonic coefficients are identified from cross-sectional variation in amenity access conditional on all other attributes; unless the listing premium varies systematically with amenity proximity (conditional on price level, location, and structural characteristics), the coefficients are unbiased. Second, the spatial lag partially absorbs any neighborhood-level variation in listing strategy. Third, our quantile regression approach is more robust to distributional misspecification than OLS (Koenker, 2005), which is relevant if the listing-to-sale ratio varies across the price distribution.

7. Discussion

7.1. Heterogeneity as the Rule, Not the Exception

Our central finding is that amenity capitalization is heterogeneous across the price distribution. The omnibus Wald test for houses ($\chi^2(7) = 35.74$, $p < 0.001$) rejects the OLS homogeneity assumption. This result aligns with the international literature (Zietz et al., 2008; Liao & Wang, 2012; Mak et al., 2010) but extends it to the Quebec context for the first time.

The economic interpretation is straightforward: households at different points in the price distribution have different preferences, constraints, and substitution possibilities. Transit-dependent households purchasing affordable homes value transit proximity three times more than affluent households with private vehicles. Conversely, luxury homebuyers may value the prestige and

green space of golf course neighborhoods more than their budgetary significance would suggest. These patterns are consistent with the sorting equilibrium framework of [Bayer et al. \(2007\)](#), where heterogeneous households sort across locations based on their willingness to pay for local amenities, generating the distributional heterogeneity we document.

The practical significance of this heterogeneity is considerable. The OLS estimate of the highway effect (0.133) overestimates the effect for luxury homes by a factor of 4 (0.133 vs. 0.034) and underestimates it for affordable homes by 35% (0.133 vs. 0.205). Any policy analysis based on the OLS coefficient would systematically misrepresent the welfare implications across income groups.

7.2. The Hospital Proximity Puzzle

The sign reversal of hospital proximity across quantiles—negative for cheap homes, positive for expensive ones—is an original finding that merits discussion. We propose three complementary explanations:

1. **Institutional disamenity vs. neighborhood quality.** Affordable homes near hospitals may be in dense, noisy, institutional corridors (ambulance traffic, surface parking). Luxury homes near hospitals may be in established affluent neighborhoods where major medical centers have located historically. In Quebec, major hospitals (McGill University Health Centre, Hôpital de Montréal pour Enfants, Centre hospitalier de l'Université Laval) are often situated in prestigious neighborhoods, creating a positive correlation between hospital proximity and neighborhood wealth at the upper end of the distribution.
2. **Sorting.** High-income households may sort into hospital-proximate neighborhoods for accessibility to specialized medical services, while low-income households near hospitals occupy lower-quality adjacent housing stock. [Bayer et al. \(2007\)](#) document this type of amenity-driven sorting in U.S. metropolitan areas.
3. **Omitted variable bias.** Hospital proximity may proxy for neighborhood density and urbanization, which have heterogeneous effects across the price distribution. The spatial lag partially controls for this, as the hospital coefficient becomes insignificant in the 2SQR median model ($\beta = -0.006$, $p > 0.10$), suggesting that the OLS hospital effect is substantially confounded by spatial spillovers.

7.3. Cross-Segment Patterns

The four segments reveal distinct capitalization patterns that reflect fundamental differences in buyer populations and product characteristics:

Houses show the richest heterogeneity, with 4/6 amenities significantly varying across quantiles. The structural model is well-specified ($R^2 = 0.62$) and amenity effects are precisely estimated. The house market spans the widest price range (10:1 ratio between $\tau = 0.90$ and $\tau = 0.10$), providing

the most statistical power to detect distributional heterogeneity.

Condos show no significant heterogeneity, consistent with a more homogeneous urban product where amenity access is relatively uniform. The high R^2 (0.67) suggests that structural variables (floor area, bathrooms) dominate price determination. The predominantly urban sample (62.4% Montreal) limits geographic variation in amenity access, reducing the power of inter-quantile tests.

Plex shows intermediate heterogeneity, with investment logic adding a layer of complexity: lower-priced plex buyers may be first-time investors prioritizing accessibility (for tenant convenience), while higher-priced plex buyers focus on tenant quality neighborhoods. The dual owner-occupant/investor nature of the plex market creates a natural source of preference heterogeneity (Pryce & Oates, 2008).

Land shows the weakest structural model ($R^2 = 0.15$) but the strongest spatial dependence ($\rho = 0.32$), consistent with the importance of comparable-sales and neighborhood effects in land valuation where the physical product is undifferentiated (Colwell & Munneke, 1997). The significant heterogeneity in school and golf coefficients may reflect different buyer populations: residential builders (sensitive to school quality and family-friendliness) versus recreational/agricultural buyers (sensitive to golf and rural amenities).

7.4. Spatial Spillovers

The spatial lag ρ is significant in the house segment across all quantiles (0.22–0.28), while the condo and land segments show weaker spatial dependence ($\rho \approx 0.04$ –0.05 at the median). This implies that a 10% increase in average neighboring prices is associated with a 0.4–2.8% increase in a property's price, depending on segment and quantile. The approximate indirect effects decomposition suggests that 4–23% of the total amenity effect operates through spatial spillovers rather than direct capitalization, with the share being largest for houses.

The variation of ρ across quantiles is itself informative. For condos, the spatial lag declines from 0.34 to zero as we move from cheap to luxury properties, suggesting that luxury condo prices are determined more by unique unit characteristics (floor level, views, finishes) than by neighborhood comparables. This is consistent with the thin-market hypothesis: luxury condos are sufficiently differentiated that neighboring prices provide limited information for valuation (Haurin, 1988).

For plex, the opposite pattern (ρ increasing with quantile) suggests that higher-valued investment properties are more sensitive to comparable-sales evidence. Professional investors purchasing expensive multiplexes likely rely more heavily on comparable transaction data, creating stronger spatial price dependencies (Shimizu et al., 2010).

7.5. Policy Implications

While these results reflect conditional associations rather than causal effects, several policy-relevant observations emerge from the heterogeneity patterns:

1. **Property taxation.** Assessments based on mean hedonic estimates will systematically over-value amenity effects for expensive properties (e.g., highway) and under-value them for affordable ones (or vice versa, depending on the amenity). Municipal assessment offices in Quebec use computer-assisted mass appraisal (CAMA) models that are typically OLS-based; incorporating distributional heterogeneity could improve assessment equity across property classes.
2. **Urban planning.** Transit investments have the largest implicit price for affordable housing—a progressive finding that supports transit-oriented development in lower-income neighborhoods. Our estimates suggest that a new transit connection would increase the value of a \$275,000 home by approximately \$53,567 (19.5%), compared to \$80,978 (5.4%) for a \$1.49M home. In proportional terms, the transit premium is nearly four times larger for affordable properties, supporting the case for transit-oriented affordable housing policy ([Renne, 2016](#)).
3. **Cost-benefit analysis.** Park investments capitalize more into affordable homes (\$35,648 at $\tau = 0.10$) than luxury homes in absolute terms, and substantially more in proportional terms (13.0% vs. 1.4%). This suggests that park development in affordable neighborhoods generates more housing value per dollar of public investment, providing a fiscal rationale for equitable park distribution ([Wolch et al., 2014](#)).
4. **Environmental justice.** The school proximity discount is amplified for luxury homes in absolute dollar terms (\$156,225 vs. \$26,608), suggesting that wealthy households are more able to sort away from school-proximate (i.e., dense, high-traffic) locations. This sorting reinforces spatial income inequality and may contribute to school quality disparities through property tax mechanisms ([Tiebout, 1956](#)).
5. **Mortgage risk assessment.** Lenders and insurers (including the Canada Mortgage and Housing Corporation, CMHC) use automated valuation models for underwriting. Our findings suggest that AVM accuracy could be improved by incorporating quantile-dependent amenity valuations, particularly for properties at the tails of the price distribution where OLS-based models perform worst.

7.6. Limitations and Caveats

The following limitations circumscribe the interpretation of our results and define priorities for future work:

1. **Listing prices.** We use listing prices rather than transaction prices. Listing prices may reflect strategic pricing behavior; the listing-to-sale ratio may vary by segment and quantile (luxury

listings may be more aspirationally priced), and amenity proximity may correlate with bargaining power. Our coefficients are valid if the listing premium is conditionally independent of amenity access, but this assumption cannot be tested without matched transaction data.

2. **Cross-sectional design.** The analysis cannot identify causal effects. Endogeneity from household sorting (Bayer et al., 2007) and unobserved neighborhood quality remains a concern, though the spatial lag and regional fixed effects partially address correlated omitted variables. The ideal design would exploit quasi-experimental variation in amenity provision (e.g., new park openings, transit extensions), which requires panel data.
3. **Binary amenity indicators.** The Centris “Amenities Nearby” field depends on broker classification, introducing measurement error and broker discretion. The variables are binary (near/far) without distance gradients, which attenuates estimated effects and prevents the identification of distance-decay patterns (Crompton, 2001). “Near school” does not capture school quality; “near transit” does not distinguish bus, metro, REM, or commuter rail; “near hospital” does not distinguish major academic hospitals from small clinics.
4. **Sample selection.** The condo retention rate (11%) creates possible selection bias if listings with missing floor area are systematically different from retained observations.
5. **Spatial weights specification.** The K -nearest-neighbor weights with $K = 10$ are constructed within each segment. Neighbors can only belong to the same property type. The results are qualitatively robust to alternative K values (Section 6), but we have not tested distance-band weights or cross-segment neighbor definitions.
6. **Residual spatial autocorrelation.** Moran’s I on 2SQR residuals remains substantial (0.35–0.40), indicating that the spatial lag does not fully absorb spatial structure. Unobserved neighborhood characteristics (crime, aesthetics, school quality) likely generate persistent spatial dependence.
7. **Generated regressor problem.** The 2SQR estimator uses first-stage predicted values as if they were observed, which does not affect consistency but may bias standard errors. Our bootstrap procedure partially addresses this, but it resamples only the second stage.
8. **Bootstrap and spatial dependence.** The paired bootstrap resamples observations independently, breaking the spatial structure of the data. Block bootstrap or subsampling approaches would provide more conservative inference (Lahiri, 2003) but are computationally expensive and require assumptions about block size.
9. **Approximate effect decomposition.** The direct and indirect effects reported use scalar approximations rather than the formal LeSage–Pace matrix-based decomposition. As discussed in Section 3, the approximation error is small for our specifications, but formal computation would strengthen the results.
10. **Scope of interpretation.** The estimated coefficients are quantile-specific capitalization gradients—conditional associations in listing prices—not structural demand parameters or causal welfare

measures. Policy implications should be interpreted accordingly.

8. Conclusion

This paper combines three dimensions of housing market analysis that are usually studied separately: quantile-specific hedonic pricing, spatial dependence through an instrumented spatial lag, and cross-segment comparison across houses, condominiums, multiplexes, and vacant land in the Quebec residential market. Using 42,481 Centris listings, we document substantial heterogeneity in amenity capitalization that standard OLS hedonic models mask. The omnibus inter-quantile test decisively rejects coefficient homogeneity for the house segment. Highway, hospital, park, and transit proximity are each associated with significantly different capitalization gradients at the 10th versus 90th percentiles, with hospital proximity exhibiting a sign reversal from negative to positive. Spatial spillovers are significant across all segments, with approximate indirect effects accounting for 22–32% of total capitalization.

These findings suggest that reporting a single mean implicit price for each housing attribute can mislead both researchers and policymakers. The OLS highway coefficient overstates the association for luxury homes by a factor of four while understating it for affordable homes by 35%. For hospital proximity, the mean estimate (−4.9%) averages a negative association for lower-priced homes with a positive association for higher-priced homes, and is therefore not representative of either tail. The policy implication is that mean-based hedonic estimates may misrepresent the distributional incidence of amenity investments, property tax assessments, and cost-benefit analyses.

We emphasize that the estimated coefficients are quantile-specific capitalization gradients—spatially adjusted conditional associations in listing prices—rather than causal effects of amenity access. The cross-sectional design, reliance on listing prices, use of broker-classified binary amenity indicators, and residual spatial autocorrelation in the 2SQR model all limit the causal interpretation. These caveats notwithstanding, the consistency of the heterogeneity patterns across multiple amenities and segments, and their alignment with theoretical predictions about preference heterogeneity, strengthens confidence in the core finding.

Several avenues for future research would address the current limitations and deepen the analysis:

1. **Transaction-price replication.** Reproducing the analysis with sale prices from Quebec's *Registre foncier* would validate the listing-price results and quantify any listing-premium bias.
2. **Continuous GIS distances.** Constructing distance measures to each amenity type would enable estimation of distance-decay functions and reduce the attenuation bias from binary indicators.
3. **Amenity quality measures.** Incorporating school quality scores, transit frequency data, hospital classification, and park size would separate amenity quality from mere proximity.
4. **Panel and quasi-experimental designs.** Difference-in-differences or event study designs

around new transit openings, park constructions, or hospital relocations would strengthen causal identification ([Kuminoff et al., 2010](#)).

5. **Geographically weighted quantile regression.** GWQR models ([Chen et al., 2012](#)) would reveal whether distributional heterogeneity itself varies across space—for example, whether the highway gradient differs between urban and rural Quebec.
6. **Machine learning benchmarks.** Quantile random forests ([Meinshausen, 2006](#)) or gradient boosted quantile regression ([Friedman, 2001](#)) could serve as flexible prediction benchmarks, though at the cost of coefficient interpretability.
7. **Spatial Durbin quantile models.** Including spatially lagged exogenous variables directly in the quantile regression would test whether neighbor characteristics have direct effects beyond the spatial lag.
8. **Temporal dynamics.** Multi-period analysis would reveal whether distributional heterogeneity fluctuates with housing market cycles or represents a structural feature of the Quebec market.

The central implication of this study is that amenity capitalization in housing markets is not a single average effect, but a spatially embedded and distributionally heterogeneous process. Hedonic analyses that ignore this heterogeneity risk producing estimates that are correct on average but wrong for every subpopulation they purport to describe.

References

- Aiken, L.S., & West, S.G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Sage Publications.
- Allen, J., Clark, R., & Houde, J.F. (2014). The effect of mergers in search markets: Evidence from the Canadian mortgage industry. *American Economic Review*, 104(10), 3365–3396.
- Anderson, S.T., & West, S.E. (2006). Open space, residential property values, and spatial context. *Regional Science and Urban Economics*, 36(6), 773–789.
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers.
- Anselin, L. (2010). Thirty years of spatial econometrics. *Papers in Regional Science*, 89(1), 3–25.
- Bajari, P., & Benkard, C.L. (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of Political Economy*, 113(6), 1239–1276.
- Bayer, P., Ferreira, F., & McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4), 588–638.
- Belsley, D.A., Kuh, E., & Welsch, R.E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons.
- Bentley, J.L. (1975). Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9), 509–517.
- Black, S.E. (1999). Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics*, 114(2), 577–599.
- Bourassa, S.C., Hoesli, M., & Peng, V.S. (2003). Do housing submarkets really matter? *Journal of Housing Economics*, 12(1), 12–28.
- Bourassa, S.C., Cantoni, E., & Hoesli, M. (2010). Predicting house prices with spatial dependence: A comparison of alternative methods. *Journal of Real Estate Research*, 32(2), 139–159.
- Buchinsky, M. (1995). Estimating the asymptotic covariance matrix for quantile regression models: A Monte Carlo study. *Journal of Econometrics*, 68(2), 303–338.
- Buchinsky, M. (1998). Recent advances in quantile regression models: A practical guideline for empirical research. *Journal of Human Resources*, 33(1), 88–126.

- Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, 22(3), 453–474.
- Cassel, E., & Mendelsohn, R. (1985). The choice of functional forms for hedonic price equations: Comment. *Journal of Urban Economics*, 18(2), 135–142.
- Chen, Y., Liu, T., Xie, X., & Maréchal, B.G. (2020). Mapping the urban green space from a localized hedonic perspective: A quantile regression approach. *Land Use Policy*, 94, 104524.
- Chen, V.Y., Deng, W.S., Yang, T.C., & Matthews, S.A. (2012). Geographically weighted quantile regression (GWQR): An application to U.S. mortality data. *Geographical Analysis*, 44(2), 134–150.
- Chen, Y., Ye, Y., Liu, X., Yin, C., & Jones, C.A. (2025). Examining the nonlinear and spatial heterogeneity of housing prices in urban Beijing: An application of GeoShapley. *Habitat International*, 159, 103283.
- Chernozhukov, V., & Hansen, C. (2006). Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics*, 132(2), 491–525.
- Chin, T.L., & Chau, K.W. (2003). A critical review of literature on the hedonic price model. *International Journal for Housing Science and Its Applications*, 27(2), 145–165.
- Cliff, A.D., & Ord, J.K. (1981). *Spatial Processes: Models and Applications*. Pion Limited.
- Canada Mortgage and Housing Corporation (2023). *Housing Market Outlook: Canada Edition*. Ottawa: CMHC.
- Colwell, P.F., & Munneke, H.J. (1997). The structure of urban land prices. *Journal of Urban Economics*, 41(3), 321–336.
- Cribari-Neto, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics & Data Analysis*, 45(2), 215–233.
- Crompton, J.L. (2001). The impact of parks on property values: A review of the empirical evidence. *Journal of Leisure Research*, 33(1), 1–31.
- Cropper, M.L., Deck, L.B., & McConnell, K.E. (1988). On the choice of functional form for hedonic price functions. *Review of Economics and Statistics*, 70(4), 668–675.
- Debrezion, G., Pels, E., & Rietveld, P. (2007). The impact of railway stations on residential and commercial property value: A meta-analysis. *Journal of Real Estate Finance and Economics*, 35(2), 161–180.

- Des Rosiers, F., Thériault, M., & Villeneuve, P. (2011). Sorting out access and neighbourhood factors in hedonic price modelling. *Journal of Property Investment & Finance*, 18(3), 291–315.
- Des Rosiers, F., Dubé, J., & Thériault, M. (2014). Environmental contamination and housing values: A quantile regression approach. *Journal of Housing and the Built Environment*, 29(3), 475–499.
- Do, A.Q., & Grudnitski, G. (1995). Golf courses and residential house prices: An empirical examination. *Journal of Real Estate Finance and Economics*, 10(3), 261–270.
- Dubé, J., Des Rosiers, F., Thériault, M., & Dib, P. (2011). Economic impact of a supply change in mass transit in urban areas: A Canadian example. *Transportation Research Part A: Policy and Practice*, 45(1), 46–62.
- Dubé, J., Legros, D., Thériault, M., & Des Rosiers, F. (2018). Spatial dependence in the residential Canadian housing market. *Journal of Real Estate Finance and Economics*, 56(2), 223–263.
- Dubé, J., Legros, D., Thériault, M., & Des Rosiers, F. (2023). YIMBY vs. NIMBY: Impact of new social housing on house prices in Québec City. *Housing Policy Debate*, 33(4), 865–890.
- Dubé, J., Des Rosiers, F., Thériault, M., & Dib, P. (2025). Is there an additional price premium for houses exposed to urban parks? Evidence from Québec City. *Journal of Environmental Management*, 380, 124787.
- Efron, B., & Tibshirani, R.J. (1994). *An Introduction to the Bootstrap*. Chapman & Hall/CRC.
- Ekeland, I., Heckman, J.J., & Nesheim, L. (2004). Identification and estimation of hedonic models. *Journal of Political Economy*, 112(S1), S60–S109.
- Elhorst, J.P. (2010). Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis*, 5(1), 9–28.
- Farmer, S., & Lipscomb, C.A. (2010). Using quantile regression in hedonic analysis to reveal submarket competition. *Journal of Real Estate Research*, 32(4), 435–460.
- Freedman, D.A. (1981). Bootstrapping regression models. *The Annals of Statistics*, 9(6), 1218–1228.
- Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232.
- Getis, A. (2004). Spatial weights matrices. *Geographical Analysis*, 41(4), 404–410.
- Gibbons, S., & Machin, S. (2008). Valuing school quality, better transport, and lower crime: Evidence from house prices. *Oxford Review of Economic Policy*, 24(1), 99–119.

- Gibbons, S., Overman, H.G., & Patacchini, E. (2015). Spatial methods. In G. Duranton, J.V. Henderson, & W.C. Strange (Eds.), *Handbook of Regional and Urban Economics*, Vol. 5 (pp. 115–168). Elsevier.
- Glaeser, E.L., Kahn, M.E., & Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics*, 63(1), 1–24.
- Goodman, A.C., & Thibodeau, T.G. (1998). Housing market segmentation. *Journal of Housing Economics*, 7(2), 121–143.
- Gorjian, S. (2025). Integrating machine learning and hedonic regression for housing price prediction: A systematic international review. *MPRA Paper No. 125676*.
- Grudnitski, G., & Do, A.Q. (2002). Adjusting for the effects of a golf course. *The Appraisal Journal*, 70(2), 141–147.
- Gyger, T., Hauri, S., Bühlmann, S., Lehner, M., Schlesinger, J., & Sigrist, F. (2025). Explainable spatial machine learning for hedonic real estate modeling. *Real Estate Economics*, forthcoming.
- Hahn, J. (1995). Bootstrapping quantile regression estimators. *Econometric Theory*, 11(1), 105–121.
- Halvorsen, R., & Pollakowski, H.O. (1981). Choice of functional form for hedonic price equations. *Journal of Urban Economics*, 10(1), 37–49.
- Hao, L., & Naiman, D.Q. (2007). *Quantile Regression*. Sage Publications.
- Haurin, D. (1988). The duration of marketing time of residential housing. *AREUEA Journal*, 16(4), 396–410.
- Herath, S., & Maier, G. (2025). Reflections on hedonic price modeling. *Journal of European Real Estate Research*, 18(1), 5–23.
- Diao, M., & Ferreira, J. (2023). Hospital accessibility impacts on housing prices: Heterogeneity across space and price quantiles. *Journal of Housing and the Built Environment*, 38, 1875–1901.
- Theebe, M.A.J. (2004). Planes, trains, and automobiles: The impact of traffic noise on house prices. *Journal of Real Estate Finance and Economics*, 28(2–3), 209–234.
- Jin, F., & Su, L. (2001). A simple bootstrap test for spatial dependence. *Economics Letters*, 71(3), 377–383.

- Kelejian, H.H., & Prucha, I.R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, 17(1), 99–121.
- Kim, T.H., & Muller, C. (2004). Two-stage quantile regression when the first stage is based on quantile regression. *The Econometrics Journal*, 7(1), 218–231.
- Knight, J.R. (2002). Listing price, time on market, and ultimate selling price: Causes and effects of listing price changes. *Real Estate Economics*, 30(2), 213–237.
- Koenker, R., & Bassett, G.W. (1978). Regression quantiles. *Econometrica*, 46(1), 33–50.
- Koenker, R., & Bassett, G.W. (1982). Robust tests for heteroscedasticity based on regression quantiles. *Econometrica*, 50(1), 43–61.
- Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.
- Koenker, R., & Machado, J.A.F. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, 94(448), 1296–1310.
- Kostov, P. (2009). A spatial quantile regression hedonic model of agricultural land prices. *Spatial Economic Analysis*, 4(1), 53–72.
- Kuminoff, N.V., 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*, 60(3), 145–160.
- Kutner, M.H., Nachtsheim, C.J., Neter, J., & Li, W. (2004). *Applied Linear Statistical Models* (5th ed.). McGraw-Hill/Irwin.
- Lahiri, S.N. (2003). *Resampling Methods for Dependent Data*. Springer.
- LeSage, J.P., & Pace, R.K. (2009). *Introduction to Spatial Econometrics*. CRC Press.
- Liao, W.C., & Wang, X. (2012). Hedonic house prices and spatial quantile regression. *Journal of Housing Economics*, 21(1), 16–27.
- Long, J.S., & Ervin, L.H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, 54(3), 217–224.
- MacKinnon, J.G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29(3), 305–325.

- Mak, S., Choy, L., & Ho, W. (2010). Quantile regression estimates of Hong Kong real estate prices. *Urban Studies*, 47(11), 2461–2472.
- Malpezzi, S. (2003). Hedonic pricing models: A selective and applied review. In T. O'Sullivan & K. Gibb (Eds.), *Housing Economics and Public Policy* (pp. 67–89). Blackwell.
- Marquardt, D.W. (1980). Comment: You should standardize the predictor variables in your regression models. *Journal of the American Statistical Association*, 75(369), 87–91.
- McMillen, D.P. (2008). Changes in the distribution of house prices over time: Structural characteristics, neighborhood, or coefficients? *Journal of Urban Economics*, 64(3), 573–589.
- Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research*, 7, 983–999.
- Mohammad, S.I., Graham, D.J., Melo, P.C., & Anderson, R.J. (2013). A meta-analysis of the impact of rail projects on land and property values. *Transportation Research Part A: Policy and Practice*, 50, 158–170.
- Perreault, L., Bhatt, M., & Bhatt, J. (2024). Evaluating the effects of supervised consumption sites on housing prices in Montreal using hedonic models. *Drug and Alcohol Dependence Reports*, 12, 100279.
- Nelson, J.P. (2004). Meta-analysis of airport noise and hedonic property values: Problems and prospects. *Journal of Transport Economics and Policy*, 38(1), 1–27.
- Rehman, A., & Pennington-Cross, A. (2025). Spatial effect of airport proximity on house prices: A quantile regression approach. *Applied Economics*, forthcoming.
- O'Brien, R.M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673–690.
- Pace, R.K., & LeSage, J.P. (2004). Chebyshev approximation of log-determinants of spatial weight matrices. *Computational Statistics & Data Analysis*, 45(2), 179–196.
- Palm, R. (1978). Spatial segmentation of the urban housing market. *Economic Geography*, 54(3), 210–221.
- Parmeter, C.F., & Pope, J.C. (2013). Quasi-experiments and hedonic property value methods. In J.A. List & M.K. Price (Eds.), *Handbook on Experimental Economics and the Environment* (pp. 3–66). Edward Elgar.

- Pryce, G., & Oates, S. (2008). Rhetoric in the language of real estate marketing. *Housing Studies*, 23(2), 279–298.
- Renne, J.L. (2016). Transit-oriented development and affordable housing. In E. L. Birch & S. M. Wachter (Eds.), *The City in the Twenty-First Century*. University of Pennsylvania Press.
- Rey-Blanco, D., Arbués, P., López, F.A., & Páez, A. (2024). Using machine learning to identify spatial market segments: A reproducible study of housing submarkets in Madrid. *Environment and Planning B: Urban Analytics and City Science*, 51(6), 1472–1489.
- Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34–55.
- Salvati, N., Tzavidis, N., Pratesi, M., & Chambers, R. (2023). Spatial semiparametric M-quantile regression for hedonic price modelling. *AStA Advances in Statistical Analysis*, 107, 159–183.
- Shimizu, C., Nishimura, K.G., & Watanabe, T. (2010). Housing prices in Tokyo: A comparison of hedonic and repeat sales measures. *Journal of Economics and Statistics*, 230(6), 792–813.
- Sirmans, G.S., Macpherson, D.A., & Zietz, E.N. (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature*, 13(1), 1–44.
- Stakhovych, S., & Bijmolt, T.H.A. (2009). Specification of spatial models: A simulation study on weights matrices. *Papers in Regional Science*, 88(2), 389–408.
- Statistics Canada (2023). *New Housing Price Index*. Table 18-10-0205-01. Ottawa: Statistics Canada.
- Stock, J.H., & Yogo, M. (2002). Testing for weak instruments in linear IV regression. *NBER Technical Working Paper No. 284*.
- Su, L., & Yang, Z. (2007). Instrumental variable quantile estimation of spatial autoregressive models. *Working Paper, Singapore Management University*.
- Tiebout, C.M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64(5), 416–424.
- Troy, A., & Grove, J.M. (2008). Property values, parks, and crime: A hedonic analysis in Baltimore, MD. *Landscape and Urban Planning*, 87(3), 233–245.
- Waddell, P., Berry, B.J.L., & Hoch, I. (1993). Residential property values in a multinodal urban area: New evidence on the implicit price of location. *Journal of Real Estate Finance and Economics*, 7(2), 117–141.

- Watkins, C.A. (2001). The definition and identification of housing submarkets. *Environment and Planning A*, 33(12), 2235–2253.
- Wen, H., Xiao, Y., & Zhang, L. (2017). Spatial effect of river landscape on housing price: An empirical study on the Grand Canal in Hangzhou, China. *Habitat International*, 63, 34–44.
- Wolch, J.R., Byrne, J., & Newell, J.P. (2014). Urban green space, public health, and environmental justice: The challenge of making cities “just green enough.” *Landscape and Urban Planning*, 125, 234–244.
- Zietz, J., Zietz, E.N., & Sirmans, G.S. (2008). Determinants of house prices: A quantile regression approach. *Journal of Real Estate Finance and Economics*, 37(4), 317–333.

Figures

Figure not available in this build
figures/fig1_coef_plots_amenities.png

Figure 1. Amenity Coefficients by Quantile and Market Segment

Notes: Each panel shows quantile regression coefficients with 95% bootstrap confidence intervals (shaded). Dotted horizontal lines indicate OLS estimates. Declining curves (highway, transit, park) indicate stronger capitalization in affordable homes; the hospital panel shows a sign reversal from negative (affordable) to positive (luxury).

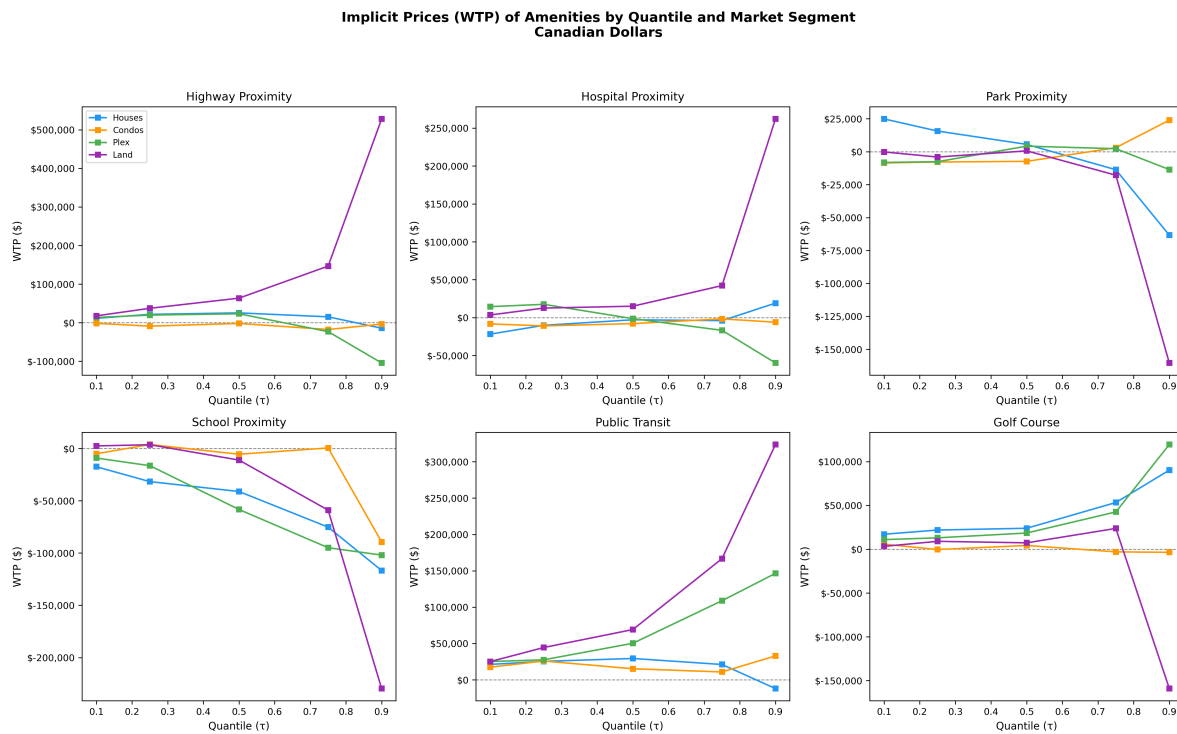


Figure 2. Implicit Prices (Willingness to Pay) by Quantile and Market Segment

Notes: WTP is computed as $P(\tau) \times [\exp(\hat{\beta}(\tau)) - 1]$, where $P(\tau)$ is the sample price at quantile τ , in Canadian dollars. The divergence between quantile-specific WTP and the OLS-implied WTP (which would be a constant line) illustrates the policy-relevant heterogeneity documented in this paper.

Figure not available in this build
 figures/fig3_ols_vs_qr_vs_sqr.png

Figure 3. Comparison of Amenity Coefficients Across Estimation Methods

Notes: Comparison of OLS, quantile regression at the median ($\tau = 0.50$), and two-stage spatial quantile regression (2SQR, $\tau = 0.50$). The spatial lag absorbs part of the amenity effect, reducing coefficient magnitudes in the 2SQR model. The attenuation is most pronounced for highway and transit proximity, suggesting that these amenity effects are substantially confounded by spatial spillovers in OLS specifications.

Figure not available in this build
 figures/fig4_pseudo_r2.png

Figure 4. Pseudo- R^2 by Quantile and Market Segment

Notes: Model fit increases with the quantile for houses and condos, suggesting that structural and amenity variables explain more variance at the upper end of the price distribution. The land segment has the lowest pseudo- R^2 throughout, reflecting the absence of structural variables and the dominance of unobserved site-specific factors (zoning, development potential).

Figure not available in this build
 figures/fig5_spatial_lag_rho.png

Figure 5. Spatial Lag Parameter (ρ) by Quantile and Market Segment

Notes: Estimated via 2SQR. Condo ρ declines sharply from 0.34 to zero as price increases, consistent with the thin-market hypothesis for luxury condos. Plex ρ increases with price, reflecting stronger comparable-sales effects for higher-valued investment properties. House and land ρ are relatively stable across quantiles.

Figure not available in this build
figures/fig6_price_distributions.png

Figure 6. Price Distributions (Log Scale) by Market Segment

Notes: Red dashed lines indicate the five quantiles $\tau \in \{0.10, 0.25, 0.50, 0.75, 0.90\}$. The right skewness of the raw price distribution (most visible for houses and plex) confirms the appropriateness of the log transformation.

Figure not available in this build
figures/fig7_wald_heatmap.png

Figure 7. Inter-Quantile Heterogeneity of Amenity Coefficients (Wald Test Statistics)

Notes: Each cell shows the Wald statistic for testing $\beta(\tau_j) = \beta(\tau_{j+1})$ between adjacent quantiles. Warm colors indicate stronger heterogeneity. The house segment exhibits the most pervasive heterogeneity, with highway and transit showing significant shifts across multiple adjacent quantile pairs. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.