Université du Québec en Outaouais

Département des sciences administratives

Working Paper No. 5

Airbnb, Residential Rents, and Housing Market Pressure: A Hedonic and Spatial Econometric Analysis

Simon-Pierre Boucher[1]

Département des sciences administratives
Université du Québec en Outaouais

This version: May 2026

Version 1.0

[1]Professeur, Département des sciences administratives, Université du Québec en Outaouais (UQO), 283 boulevard Alexandre-Taché, Gatineau, QC J9A 1L8, Canada. Email: simon-pierre.boucher@uqo.ca.

Abstract

This paper investigates the relationship between Airbnb short-term rental activity and residential rents in Quebec, Canada. Using cross-sectional microdata comprising approximately 5,000 Airbnb listings and 8,300 residential rental listings, we employ a hedonic pricing framework augmented with spatial econometric techniques to quantify the conditional association between nearby Airbnb presence and monthly rents. For each rental listing we construct Airbnb exposure measures within 250 m, 500 m, 1 km, and 2 km buffers using Haversine distances. Our baseline estimates indicate that an additional Airbnb listing within 500 m is associated with a statistically significant increase in monthly rent of approximately 0.3–0.5%, controlling for dwelling characteristics, building type, and city fixed effects. Quantile regressions reveal that this association is stronger at the upper tail of the rent distribution. A complementary hedonic model of Airbnb nightly prices indicates that short-term rental pricing is driven primarily by listing characteristics, with no significant premium attached to city-level rents. Robustness checks—including alternative buffer radii, subsample analyses, and machine-learning benchmarks—confirm the stability of the rent–exposure association. We discuss the policy implications of these results for housing affordability and short-term rental regulation, while cautioning that cross-sectional associations should not be interpreted as causal effects.

Keywords: Airbnb, short-term rentals, housing rents, hedonic pricing, spatial econometrics, housing affordability, Quebec

JEL Classification: R21, R31, L83, C21

1. Introduction

The rapid expansion of short-term rental platforms has fundamentally transformed urban housing markets around the world. Since its founding in 2008, Airbnb has grown from a modest home-sharing service into a global hospitality platform with millions of listings across virtually every major city, an archetypal case of disruptive innovation in tourism accommodation (Guttentag 2015) and of the broader rise of peer-to-peer markets (Einav et al. 2016; Sundararajan 2016). By enabling property owners to rent their dwellings—or portions thereof—to short-term visitors at nightly rates that often exceed what a long-term tenant would pay on a monthly basis, the platform has created powerful economic incentives to divert housing units from the residential rental market. The same flexible peer supply that generates large welfare gains in the accommodation market (Zervas et al. 2017; Farronato and Fradkin 2022) is, from the housing market's perspective, a competing use for the residential stock.

This diversion mechanism sits at the heart of a growing policy concern: that the proliferation of Airbnb listings may contribute to rising residential rents and reduced housing affordability, particularly in cities with significant tourist appeal. A now-substantial empirical literature supports the concern: studies exploiting cross-market variation (Barron et al. 2021; Horn and Merante 2017), regulatory discontinuities (Koster et al. 2021; Valentin 2021; Duso et al. 2024), and within-city diffusion (Garcia-López et al. 2020; Franco and Santos 2021) consistently find that STR activity raises rents and house prices, with effects that are modest per listing but spatially concentrated. The economic logic is straightforward. When landlords can earn higher returns by listing a unit on Airbnb than by renting it to a long-term tenant, the effective supply of long-term rental housing contracts. On the demand side, Airbnb may attract additional visitors to a neighbourhood, increasing foot traffic, consumption amenities, and—through general equilibrium effects—the desirability of the area for both tourists and residents. The net effect on rents is theoretically ambiguous: supply withdrawal pushes rents upward, while the sign and magnitude of demand-side amenity effects depend on the local context, the composition of Airbnb listings, and the degree of market segmentation between short-term and long-term rentals.

The province of Quebec, and the city of Montreal in particular, provides a compelling setting in which to study these dynamics. Montreal is Canada's second-largest city and one of North America's foremost tourist destinations, attracting over 11 million visitors annually prior to the COVID-19 pandemic. Its dense, walkable neighbourhoods, vibrant cultural scene, and historic architecture make it especially attractive for short-term rental guests, and national analyses identify it as one of Canada's largest and most commercialised STR markets (Combs et al. 2020). At the same time, Quebec has experienced significant rental market pressures in recent years. Vacancy rates in the Montreal census metropolitan area have fallen to historically low levels, and median rents have risen substantially. Policymakers at the provincial and municipal levels have responded with a mix of regulatory interventions, including CITQ registration requirements for short-term rental operators and zoning restrictions in Montreal (Nieuwland and van Melik 2020). Yet the econometric evidence on the housing-market effects of Airbnb in the Quebec context remains remarkably thin: existing Canadian work is either national and descriptive (Combs et al. 2020) or focused on Toronto's political economy of displacement (Grisdale 2021). No study, to our knowledge, has estimated the Airbnb–rent relationship at the individual listing level for Quebec.

Estimating the effect of short-term rental platforms on residential rents poses significant empirical challenges. The most fundamental is endogeneity: Airbnb listings are not randomly assigned to locations. Hosts choose to list in neighbourhoods where rents—and tourist demand—are already high, creating a positive correlation between Airbnb density and rent levels that may reflect reverse causality or omitted neighbourhood characteristics rather than a genuine causal effect. In the ideal research design, one would exploit plausible exogenous variation in Airbnb supply—for example, a regulatory shock that differentially affected neighbourhoods—combined with panel data to control for time-invariant unobservables. Our data, however, are cross-sectional: they comprise a single snapshot of Airbnb and rental listings in Quebec, scraped in 2026. We therefore adopt a transparent hedonic pricing framework that quantifies the conditional association between Airbnb presence and rents, controlling for a rich set of dwelling characteristics and location fixed effects, while being explicit about the limitations of causal interpretation.

Our empirical strategy proceeds in several stages. First, we estimate a hedonic rent model in which the logarithm of monthly rent is regressed on the count of Airbnb listings within a spatial buffer (our preferred radius is 500 metres), controlling for the number of bedrooms, bathrooms, building type, and city fixed effects. This specification yields a semi-elasticity interpretation: the coefficient on Airbnb count measures the percentage change in rent associated with one additional nearby Airbnb listing, holding observable dwelling characteristics constant. Second, we estimate a complementary hedonic model of Airbnb nightly prices, in which city-level mean rents serve as a control, allowing us to examine the bidirectional pricing relationship between the two market segments. Third, we aggregate the data to the city level to examine cross-city variation in Airbnb penetration and mean rents. Fourth, we incorporate spatial lags and vary the buffer radius (250 m, 500 m, 1 km, 2 km) to assess the spatial decay of the association. Fifth, quantile regressions at the 10th, 25th, 50th, 75th, and 90th percentiles reveal how the Airbnb–rent association varies across the conditional rent distribution. Sixth, we benchmark our parametric estimates against machine-learning methods—LASSO, elastic net, random forest, and gradient boosting—to assess predictive importance and model robustness.

To preview the central result: each additional Airbnb listing within 500 metres of a rental unit is associated with approximately 0.4% higher monthly rent, conditional on dwelling characteristics, building type, and city fixed effects. The association decays monotonically with distance, is present throughout the conditional rent distribution but strongest at its upper tail, and survives spatial-econometric treatment, city-clustered inference, functional-form changes, sample trims, and leave-one-city-out exclusions. At the median rent of roughly \$1,950 per month, the point estimate corresponds to \$6–10 per additional nearby listing—strikingly close, in order of magnitude, to the 7–13 euros per listing that Duso et al. (2024) recover from Berlin's regulatory shocks, and consistent with the per-listing effects implied by the Los Angeles ordinance evidence of Koster et al. (2021). That a transparent cross-sectional design recovers magnitudes in line with the quasi-experimental literature is, we argue, informative in both directions: it lends external credibility to our estimates and out-of-sample support to the causal designs.

This paper makes three contributions. First, it provides the first granular, listing-level econometric analysis of the Airbnb–rent nexus in Quebec—and, to our knowledge, in any Canadian market—drawing on microdata with precise geographic coordinates that allow exact distance-based matching between Airbnb and rental listings. Prior Canadian evidence is descriptive (Combs et al. 2020) or qualitative (Grisdale 2021); we complement it with formal estimation. Second, it applies a multi-method approach—hedonic regressions (Rosen 1974), spatial econometrics estimated by generalised moments (Kelejian and Prucha 1998, 1999), quantile regressions (Koenker and Bassett 1978), and machine-learning benchmarks (Mullainathan and Spiess 2017)—to a single dataset, enabling direct comparison of results across methodological frameworks and, in particular, delivering the first distributional (quantile-level) characterisation of the STR–rent association for a Canadian market. Third, it contributes to the ongoing policy debate about short-term rental regulation in Canadian cities by providing empirically grounded estimates of the magnitude and spatial reach of the Airbnb–rent association—quantities that bear directly on the geographic targeting of Quebec's registration-based regime—even as it highlights the limitations inherent in cross-sectional identification.

The remainder of the paper is organised as follows. Section 2 reviews the related literature on short-term rentals and housing markets. Section 3 describes the data sources, cleaning procedures, and spatial merge strategy. Section 4 presents the econometric framework. Section 5 reports the main empirical results. Section 6 provides robustness checks and sensitivity analyses. Section 7 discusses policy implications and limitations. Section 8 concludes.

2. Literature Review

This section reviews the empirical and methodological literatures on which the paper builds. We organise the review thematically rather than chronologically: we first survey the evidence on short-term rental (STR) platforms and housing costs, moving from the early correlational studies to the more recent quasi-experimental designs (Section 2.1); we then turn to the Canadian and Quebec evidence (Section 2.2), the broader economics of peer-to-peer accommodation markets (Section 2.3), and the pricing of STR listings themselves (Section 2.4). The remaining subsections cover the methodological foundations of the paper: hedonic pricing theory and practice (Section 2.5), spatial econometrics (Section 2.6), distributional analysis via quantile regression (Section 2.7), and machine learning in housing economics (Section 2.8). Section 2.9 reviews the tourism-gentrification and regulation literature that frames the policy discussion, and Section 2.10 states the gap this paper fills.

2.1. Short-Term Rentals and Housing Costs

Housing costs respond to shifts in effective supply, and markets with inelastic supply capitalise demand shocks into prices and rents (Glaeser et al. 2006; Saiz 2010; Gyourko et al. 2013). STR platforms create precisely such a shift at the margin: by raising the return to allocating a dwelling to visitors rather than tenants, they can withdraw units from the long-term rental stock. The empirical literature testing this mechanism has developed in two waves.

Correlational and panel evidence. The seminal large-scale study is Barron et al. (2021), who exploit zipcode-level variation in Airbnb penetration across the United States, instrumenting Airbnb growth with the interaction of Google Trends search interest and a measure of touristiness. They estimate that a 1% increase in Airbnb listings raises rents by 0.018% and house prices by 0.026%, and attribute the effect to the reallocation of units from long-term to short-term supply. Horn and Merante (2017) use listing-level asking rents in Boston and find that a one-standard-deviation increase in Airbnb density is associated with rents approximately 0.4% higher, an effect concentrated where Airbnb listings are entire homes. Sheppard and Udell (2016) document property-value capitalisation in New York City, and Wachsmuth and Weisler (2018) show that the "rent gap" opened by STR revenue potential is systematically related to gentrification pressure. In a similar spirit, Ayouba et al. (2020) apply spatial panel methods to eight French cities and find significant rent effects of Airbnb density in Lyon, Montpellier, and Paris but not elsewhere—early evidence that the association is heterogeneous across markets. Using proprietary platform data and a structural model of landlord choice, Li et al. (2022) show that Airbnb entry shifts landlords at the margin between long-term and short-term supply, with the incidence falling on long-term tenants in high-demand markets.

Quasi-experimental evidence. A second wave exploits policy discontinuities and regulatory shocks. Garcia-López et al. (2020) use the staggered, spatially uneven diffusion of Airbnb across Barcelona neighbourhoods and estimate that Airbnb activity raised rents by 1.9% and transaction prices by 4.6–7% in high-activity areas. Koster et al. (2021) exploit Home-Sharing Ordinances adopted by some Los Angeles-area cities but not others in a spatial regression-discontinuity design:

ordinances reduced Airbnb listings by roughly 50% and lowered house prices and rents by around 2% in treated areas, implying a substantial positive effect of STR activity in the absence of regulation. Valentin (2021) finds that New Orleans' licensing reform capitalised into property values where STR use was legalised. Duso et al. (2024) use two Berlin regulatory shocks and estimate that each nearby Airbnb listing raises asking rents by roughly 7–13 euros per month, while Franco and Santos (2021) document positive effects of Airbnb density on both prices and rents across Portuguese municipalities, strongest in tourist-intensive locations. For London, Shabrina et al. (2022) link neighbourhood-level Airbnb "misuse" (de facto commercial operation) to higher local rents.

Taken together, this literature finds positive, modest-per-listing, spatially concentrated associations between STR activity and residential housing costs, with quasi-experimental designs generally confirming—and sometimes exceeding—the magnitudes suggested by earlier correlational work. Our estimates are best read against this benchmark: a cross-sectional design cannot adjudicate causality, but it can establish whether the conditional correlations in a previously unstudied market are consistent with the causal magnitudes established elsewhere.

2.2. Canadian and Quebec Evidence

Evidence for Canada is comparatively scarce and largely descriptive. Combs et al. (2020) provide the first national analysis of Canadian STR markets, documenting rapid but spatially uneven growth, strong revenue concentration among commercial multi-listing operators, and an estimate of roughly 31,000 housing units removed from Canadian long-term rental markets by frequently rented entire-home listings in 2018—with Montreal among the three dominant markets. Grisdale (2021) traces the political economy of STR-driven displacement in Toronto, linking commercial Airbnb operation to rental-supply loss in central neighbourhoods. Beyond descriptive market measurement and critical urban geography, however, there is little listing-level econometric evidence for Canadian markets, and none, to our knowledge, for Quebec—a striking gap given that the province combines one of Canada's largest STR markets, distinctive provincial registration requirements, and acute rental-market tightness. This paper addresses that gap.

2.3. The Economics of Peer-to-Peer Accommodation

A complementary literature studies STR platforms as peer-to-peer markets. Einav et al. (2016) analyse how platforms reduce transaction costs and enable trade in underutilised assets, and Sundararajan (2016) surveys the rise of "crowd-based capitalism." Guttentag (2015) frames Airbnb as a disruptive innovation in tourism accommodation. On market impacts, Zervas et al. (2017) estimate that Airbnb entry reduced hotel revenues in the most affected Texas segments by 8–10%, and Farronato and Fradkin (2022) quantify the welfare effects of peer entry in the accommodation industry, showing that flexible peer supply expands capacity precisely when demand peaks, generating consumer surplus while eroding hotel profits. This literature matters for our purposes because it establishes the economic incentive at the heart of the housing channel: peer hosts respond elastically to short-term rental returns, and the same responsiveness that creates accommodation-market surplus is what reallocates dwellings away from long-term tenants when STR returns exceed residential rents.

2.4. Pricing of Short-Term Rental Listings

A distinct strand estimates hedonic models of Airbnb nightly prices, which informs our Model 2. Wang and Nicolau (2017) study some 180,000 listings in 33 cities and find that entire-home status, capacity, and host attributes (notably superhost status) are dominant price determinants. Gibbs et al. (2018) estimate hedonic price equations for five Canadian Airbnb markets—the study closest to our Model 2 in geographic scope—and document significant premia for entire homes and location. Ert et al. (2016) show that host reputation and trust signals are capitalised into Airbnb prices. Notably, these studies generally find positive superhost premia, whereas our within-city estimates yield a superhost *discount*; we return to this disagreement in Sections 5 and 7 and interpret it as a composition effect specific to Quebec's mix of urban and resort listings.

2.5. Hedonic Pricing: Theory and Practice

The hedonic framework descends from Lancaster (1966), who recast consumer demand as demand for characteristics, and Rosen (1974), who formalised the market equilibrium in which the prices of differentiated products reveal implicit characteristic prices. In housing applications, the hedonic price of a locational attribute—here, exposure to nearby STR activity—is identified from cross-sectional price variation conditional on structural characteristics (Palmquist 2005). Surveys of five decades of empirical practice document both the workhorse status and the fragility of hedonic estimates: Sirmans et al. (2005) catalogue the composition of hundreds of published models; Malpezzi (2003) reviews functional-form and specification choices; and Kuminoff et al. (2010) show in large-scale simulations that spatial fixed effects and flexible functional forms are the most effective safeguards against omitted-variable bias in cross-sectional hedonics—a prescription our specification follows (city fixed effects; count, density, and logarithmic exposure forms; extensive robustness analysis). Parmeter et al. (2010) discuss the identification threats that remain.

2.6. Spatial Econometrics in Housing Research

Housing prices are spatially dependent: nearby dwellings share unobserved amenities, and pricing spillovers propagate through comparable-based valuation. Ignoring this dependence invalidates inference and can bias coefficients when spatially correlated unobservables are also correlated with regressors (Dubin 1988; Can 1992). The canonical modelling responses are the spatial autoregressive (SAR) and spatial error (SEM) models (Anselin 1988), with feasible estimation by spatial two-stage least squares and generalised moments developed by Kelejian and Prucha (1998) and Kelejian and Prucha (1999)—the estimators we employ—and comprehensive treatment in LeSage and Pace (2009). Our buffer-count exposure variable is itself a spatially weighted aggregate, so the spatial models play a dual role in this paper: they absorb residual spatial dependence, and they provide a conservative check on whether the Airbnb–rent association merely proxies for spatially correlated unobservables.

2.7. Distributional Effects and Quantile Regression

Mean regression can mask heterogeneity across the conditional price distribution. Quantile regression (Koenker and Bassett 1978; Koenker and Hallock 2001) recovers characteristic prices at arbitrary quantiles, and housing applications consistently find that implicit prices differ across the distribution: Zietz et al. (2008) show that buyers of higher-priced homes value certain attributes differently from buyers of lower-priced homes, and McMillen (2008) decomposes changes in the house-price distribution into characteristics and coefficient effects. In the STR context, distributional heterogeneity is economically meaningful: if STR pressure is concentrated where rents are already high, the affordability consequences differ sharply from those of a uniform shift. Our Model 5 provides, to our knowledge, the first quantile-level evidence on the Airbnb–rent association for a Canadian market.

2.8. Machine Learning in Housing Economics

Machine-learning methods enter modern hedonic practice in two roles: as prediction benchmarks and as data-driven specification checks (Mullainathan and Spiess 2017; Athey and Imbens 2019). We use the LASSO (Tibshirani 1996) and the elastic net (Zou and Hastie 2005) for regularised variable selection, and random forests (Breiman 2001) and gradient boosting (Friedman 2001) as flexible nonlinear benchmarks, with SHAP values (Lundberg and Lee 2017) for feature attribution; all models are implemented in scikit-learn (Pedregosa et al. 2011). The comparison between OLS and the flexible learners serves an econometric purpose articulated by Mullainathan and Spiess (2017): if a linear hedonic attains predictive accuracy close to the flexible frontier, linearity is an adequate approximation for inference; where it falls short, the gap signals nonlinearities—in our data, primarily geographic—that fixed effects and buffer construction must absorb.

2.9. Tourism Gentrification and Short-Term Rental Regulation

Finally, a critical urban literature situates STRs within longer-running processes of tourism-driven neighbourhood change. Gotham (2005) coined “tourism gentrification” to describe the transformation of New Orleans’ Vieux Carré; Cocola-Gant (2016) identifies holiday rentals as a distinct gentrification battlefront in which transient visitors replace residents; and Wachsmuth and

Weisler (2018) connect this process explicitly to platform-enabled rent gaps. The commercialisation of nominally peer-to-peer platforms is central to this critique: Ke (2017) documents that professional multi-listing hosts account for a disproportionate share of Airbnb supply and revenue. On the policy side, Gurran and Phibbs (2017) ask how urban planners should respond to Airbnb; Nieuwland and van Melik (2020) compare regulatory regimes across cities, from outright prohibition to registration and cap systems; and Agyeman et al. (2020) examine the equity dimensions of STR regulation. In Quebec, provincial law requires STR operators to register with the Corporation de l'industrie touristique du Québec (CITQ), and Montreal restricts commercial STRs to designated zones—a regime whose spatial targeting our decay results speak to directly.

2.10. Positioning and Contribution

Three gaps emerge from this review. First, despite a mature international literature spanning correlational and quasi-experimental designs, there is no listing-level econometric evidence on the STR–rent relationship for Quebec, and Canadian evidence more broadly stops at descriptive market measurement (Combs et al. 2020) and critical urban geography (Grisdale 2021). Second, few studies examine the *distributional* incidence of the association—most report mean effects—although the affordability debate turns on who bears the pressure. Third, the pricing side and the rent side of the market are rarely analysed jointly in the same data, despite the opportunity-cost logic that connects them. This paper addresses all three gaps: it provides the first listing-level, multi-method quantification of the Airbnb–rent association in Quebec, characterises the association across the full conditional rent distribution, and estimates both sides of the market with mutually consistent variable constructions. We are equally deliberate about what the paper does not do: with a single cross-section we do not identify causal effects, and we position our estimates as conditional correlations to be read against—and, as it turns out, found consistent with—the quasi-experimental magnitudes surveyed above.

3. Data

This section describes the two primary datasets, the cleaning procedures applied to each, the spatial merge strategy used to link them, and the construction of key variables.

3.1. Airbnb Listings

Our Airbnb data consist of approximately 5,000 listings scraped from the Airbnb platform for the province of Quebec, Canada. Each listing record contains the following fields: listing name, city, latitude and longitude coordinates, property type, nightly price (in Canadian dollars), star rating, number of reviews, superhost status, guest-favourite designation, pet-friendliness, and price/rating category indicators.

The listings span multiple cities across Quebec, with the largest concentrations in Montreal, Quebec City (Québec), Sherbrooke, and various resort/cottage communities in the Laurentians and Eastern Townships. Property types include *Rental unit*, *Cabin/Chalet*, *House*, *Condo*, and *Apartment*, with Rental unit being the most common category.

Cleaning. We standardise city names by merging accented and unaccented variants (e.g., "Montréal" and "Montreal") and consolidating aliases (e.g., "Quebec City" and "Québec City" mapped to "Québec"). We drop listings with missing or non-positive nightly prices, which yields a cleaned sample of 3,456 listings, and winsorise the price distribution at the 1st and 99th percentiles to mitigate the influence of extreme outliers. The log-transformed price, $\ln(\text{price\_numeric})$, is used as the dependent variable in the Airbnb pricing model. Missing ratings are imputed with the sample median, reflecting the assumption that unrated listings represent recently created properties whose quality is, on average, comparable to the centre of the distribution. Missing review counts are set to zero. An indicator variable, `is_entire_home`, is constructed to identify listings of types House, Cabin/Chalet, and Condo, which are typically rented in their entirety.

3.2. Residential Rental Listings

The rental data consist of approximately 8,300 listings scraped from Realtor.ca, Canada's primary real-estate listing platform, for the province of Quebec. Each record contains: address, latitude and longitude, property type, monthly lease rent (in Canadian dollars), building type, number of bedrooms, number of bathrooms, interior size, number of storeys, and postal code.

Building types in the rental data are dominated by *Apartment* ($n \approx 7{,}515$), followed by *House* ($n \approx 745$) and *Row/Townhouse* ($n \approx 48$). The geographic distribution is heavily concentrated in the Montreal metropolitan area, with additional clusters in Quebec City, Gatineau, Sherbrooke, and Laval.

Cleaning. Monthly rent values are extracted from the lease-rent string and converted to numeric format. We restrict the sample to listings with monthly rental periods and positive rent values. Extreme rent values are handled through winsorisation at the 1st and 99th percentiles. The log of monthly rent, $\ln(\text{rent})$, serves as the dependent variable in the hedonic rent model. Bedroom and bathroom counts are cleaned and converted to numeric types. Listings with missing geographic coordinates are dropped prior to spatial merging.

3.3. Spatial Merge Strategy

The core challenge in linking the Airbnb and rental datasets is that they share no common listing identifiers. We exploit the fact that both datasets contain precise latitude and longitude coordinates to construct a spatial buffer merge. For each rental listing i, we compute the Haversine (great-circle) distance to every Airbnb listing j and count the number of Airbnb listings falling within four concentric buffers: 250 m, 500 m, 1 km, and 2 km.

Formally, the Airbnb count for rental listing i within buffer radius r is:

$$\texttt{airbnb\_count}_{i,r} = \sum_{j=1}^{N_{\text{Airbnb}}} \mathbf{1}\left[d(i,j) \leq r\right], \tag{1}$$

where $d(i,j)$ is the Haversine distance between listings i and j, and $\mathbf{1}[\cdot]$ is the indicator function.

In addition to the count, we construct several exposure metrics within each buffer: Airbnb density (count divided by buffer area, πr^2), mean Airbnb nightly price, share of entire-home listings, mean rating, and superhost share. These variables capture not only the intensity of Airbnb activity but also its composition and quality.

At the preferred 500 m radius, the mean Airbnb count per rental listing is approximately 8 listings, with substantial variation across urban and rural areas.

3.4. City-Level Aggregation

As a complementary linkage strategy, we aggregate Airbnb data to the city level and merge city-level summary statistics—total Airbnb count, mean nightly price, share of entire-home listings—onto the rental data using the city name as the merge key. This approach sacrifices the within-city spatial variation that drives our buffer-based analysis but provides a useful benchmark for examining cross-city relationships.

3.5. Summary Statistics

Tables 1 and 2 present summary statistics for the cleaned Airbnb and rental datasets, respectively.

Table 1. Summary Statistics — Airbnb Listings

Variable	N	Mean	SD	Min	P25	Median	P75	Max
Nightly price (CAD)	3456	1233.35	2624.71	30.96	125.00	219.82	1211.76	18205.04
Rating	3456	4.80	0.22	2.25	4.75	4.86	4.94	5.00
Bedrooms	3456	2.12	1.56	0.00	1.00	2.00	3.00	25.00
Bathrooms	3456	1.30	1.09	0.00	1.00	1.00	2.00	24.00
Max guests	3456	5.40	3.41	1.00	2.00	4.00	7.00	16.00
Amenities count	3456	45.78	15.74	7.00	35.00	46.00	57.00	109.00
Number of images	3456	28.20	19.17	0.00	15.00	24.00	36.00	199.00
Number of reviews	3456	53.40	74.80	0.00	6.00	26.00	71.00	986.00
Quality score	3456	96.82	38.79	0.00	95.60	108.40	121.10	139.80

Key features of the data are worth noting. The median Airbnb nightly price is approximately

Table 2. Summary Statistics — Residential Rental Listings

Variable	N	Mean	SD	Min	P25	Median	P75	Max
Monthly rent (CAD)	8303	2141.81	764.87	995.08	1615.00	1950.00	2495.00	5000.00
Bedrooms	7925	2.05	0.94	1.00	1.00	2.00	3.00	10.00
Bathrooms	8303	1.22	0.47	0.00	1.00	1.00	1.00	8.00

$220 CAD, while the median monthly rent is approximately $1,950 CAD. There is substantial variation in both prices, reflecting the heterogeneity of listings in terms of location, size, and quality. The Airbnb data span a wide range of property types, from budget-oriented rental units to luxury chalets, while the rental data are dominated by apartment units.

3.6. Geographic Distribution

Figures 1 and 2 display the geographic distribution of Airbnb and rental listings, respectively. The spatial overlap between the two datasets is concentrated in the Montreal metropolitan area and, to a lesser extent, in Quebec City and resort regions.

Figure 3 shows the distributions of Airbnb nightly prices and monthly rents (both in levels, trimmed above the 99th percentile for readability).

3.7. Data Limitations

Several limitations of the data should be acknowledged. First, the data are cross-sectional, representing a single snapshot in time. We cannot observe how Airbnb entry or exit affects rents over time, nor can we control for time-varying confounders. Second, we do not observe host identifiers, which precludes the construction of multi-listing indicators at the host level—a variable that has been important in prior studies for distinguishing professional from casual hosts. Third, the Airbnb data are scraped from the public-facing platform and may not capture all listings, particularly those that are inactive, delisted, or hidden behind search filters. Fourth, the rental data from Realtor.ca represent *asking* rents rather than *transacted* rents; to the extent that there is systematic negotiation between posted and final rents, our dependent variable may be measured with noise. Fifth, there may be spatial selection in the data: the coverage of both platforms is likely denser in urban areas, particularly Montreal, which limits the generalisability of our findings to rural and peripheral markets.

4. Methodology

This section presents the econometric models employed to quantify the association between Airbnb activity and residential rents, discusses identification, and describes the machine-learning robustness framework.

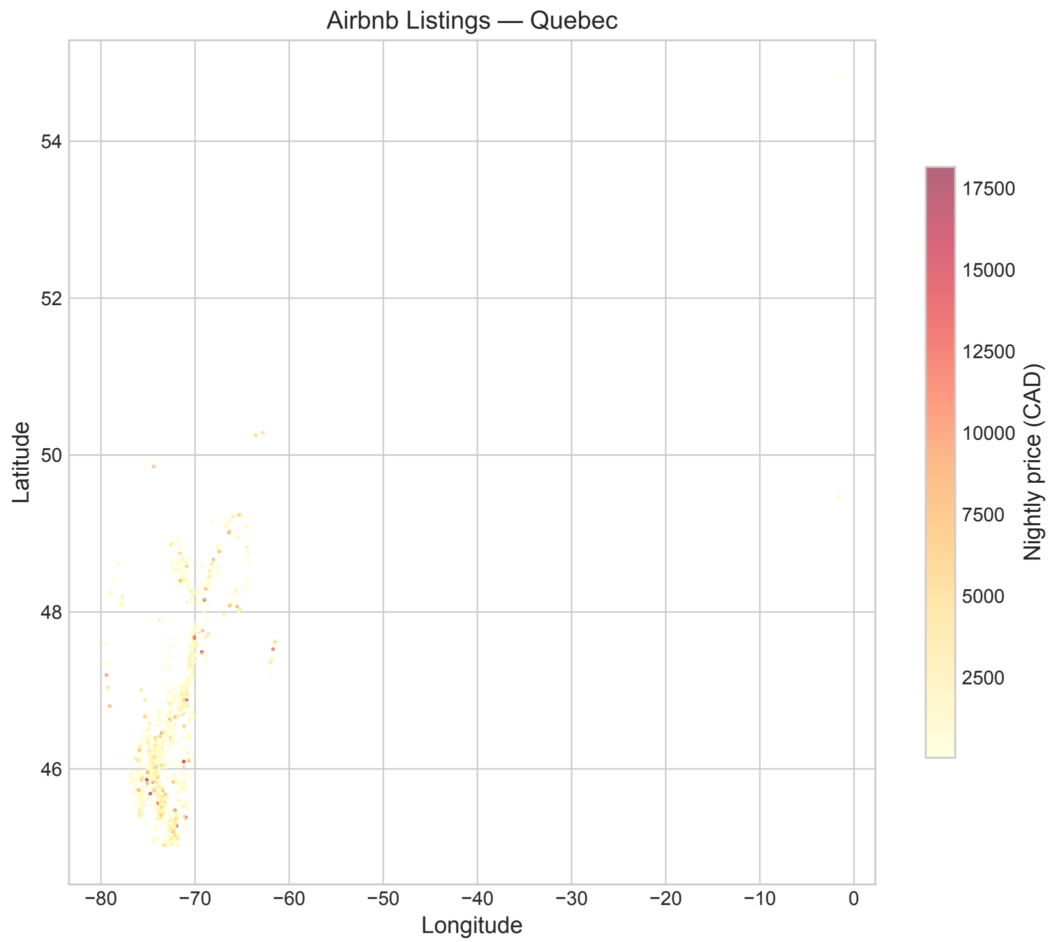


Figure 1. Geographic Distribution of Airbnb Listings in Quebec

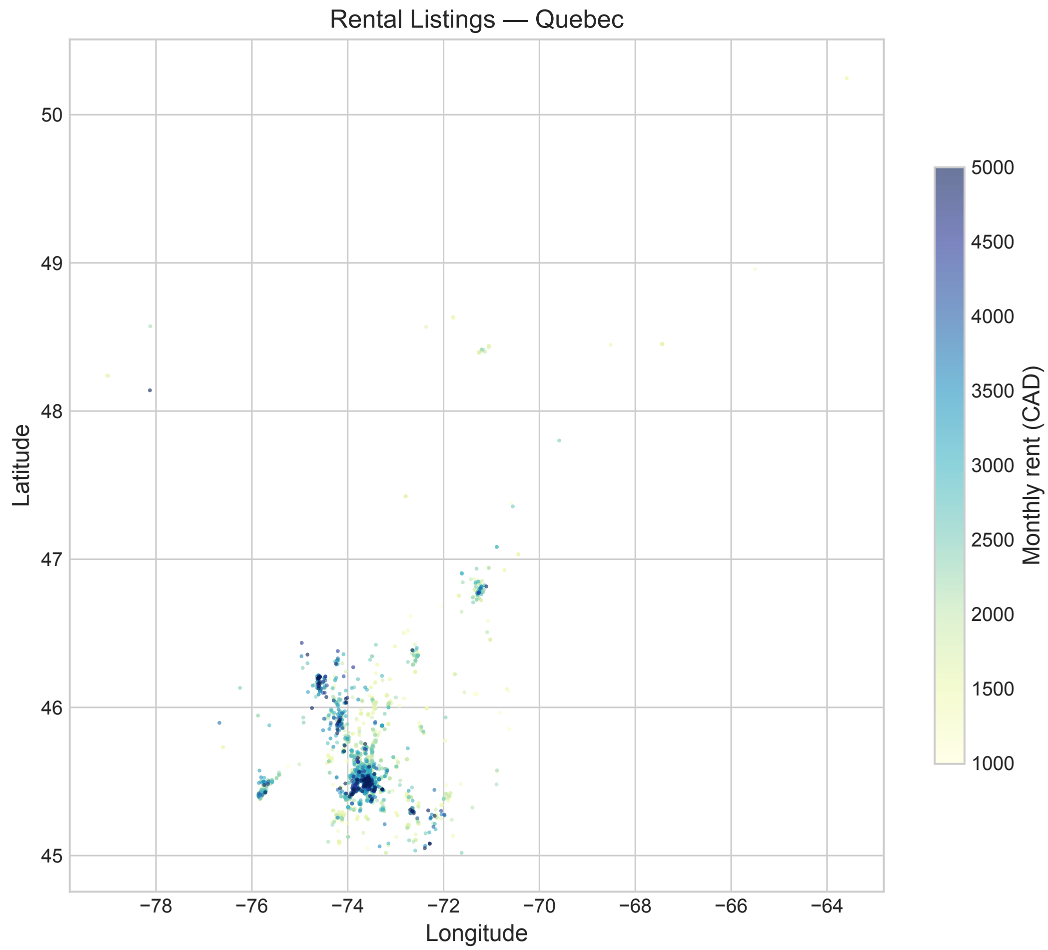
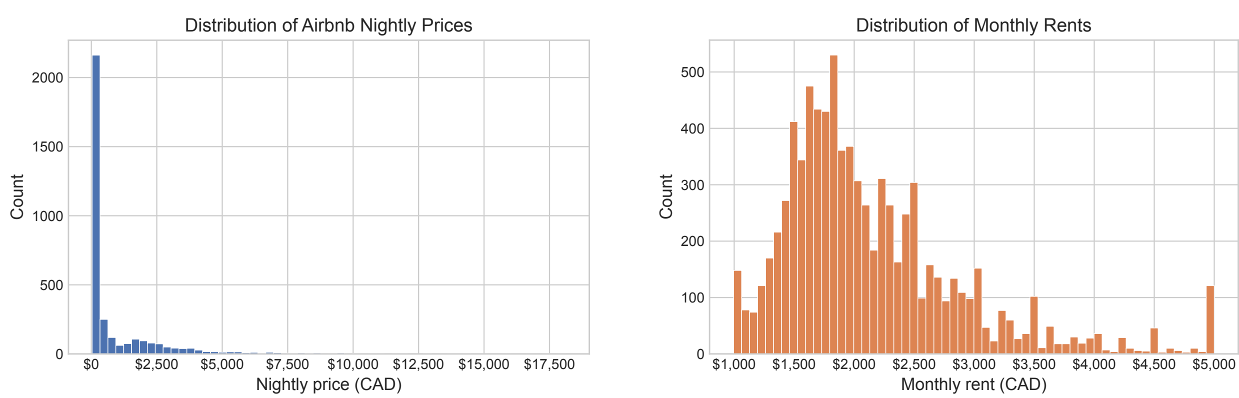


Figure 2. Geographic Distribution of Residential Rental Listings in Quebec



(a) Airbnb nightly price (CAD)

(b) Monthly rent (CAD)

Figure 3. Distributions of Airbnb Nightly Prices and Monthly Rents

4.1. Model 1: Hedonic Rent Model with Airbnb Exposure

Our baseline specification is a hedonic rent equation (Rosen 1974; Malpezzi 2003) in which the log of monthly rent is regressed on a measure of nearby Airbnb activity and a vector of dwelling-level controls:

$$\ln(\text{rent}_i) = \alpha + \beta \cdot \texttt{airbnb\_count}_{i,r} + \mathbf{X}_i'\boldsymbol{\gamma} + \sum_c \delta_c \cdot \mathbf{1}[\text{city}_i = c] + \varepsilon_i, \tag{2}$$

where rent_i is the monthly rent of listing i; $\texttt{airbnb\_count}_{i,r}$ is the number of Airbnb listings within buffer radius r of listing i (our preferred specification uses $r = 500\,\text{m}$); $\mathbf{X}_i$ is a vector of dwelling characteristics including the number of bedrooms, number of bathrooms, building type (Apartment, House, Row/Townhouse), and interior size (where available); δ_c denotes city fixed effects that absorb all time-invariant, city-level unobservables (including average neighbourhood quality, local labour-market conditions, and municipal regulations); and ε_i is a mean-zero error term.

The coefficient of interest, β, has a semi-elasticity interpretation: it measures the approximate percentage change in monthly rent associated with one additional Airbnb listing within the buffer, conditional on observed dwelling characteristics and city. A positive and statistically significant $\hat{\beta}$ is consistent with the hypothesis that Airbnb activity is associated with higher residential rents, though it does not, by itself, establish causality. The reliance on spatial fixed effects and multiple functional forms of the exposure variable follows the specification guidance of Kuminoff et al. (2010), whose simulation evidence identifies these as the most effective safeguards against omitted-variable bias in cross-sectional hedonic models.

We estimate Equation (2) by ordinary least squares (OLS) with heteroskedasticity-robust standard errors (HC1).

4.2. Model 2: Hedonic Airbnb Pricing Model

To examine the pricing determinants of Airbnb listings and the relationship between local rents and short-term rental prices, we estimate a hedonic model of Airbnb nightly prices:

$$\ln(\text{price}_j) = \alpha' + \theta \cdot \overline{\text{rent}}_c + \mathbf{Z}_j'\boldsymbol{\lambda} + \sum_c \delta_c' \cdot \mathbf{1}[\text{city}_j = c] + u_j, \tag{3}$$

where price_j is the nightly price of Airbnb listing j; $\overline{\text{rent}}_c$ is the mean log monthly rent in city c (obtained from the rental dataset); $\mathbf{Z}_j$ includes listing characteristics—the number of bedrooms and bathrooms, guest capacity, amenities count, superhost status, and the entire-home indicator; and u_j is an error term.

The coefficient θ captures the association between the local residential rent level and Airbnb pricing. A positive $\hat{\theta}$ would suggest that Airbnb hosts in higher-rent cities charge higher nightly prices, consistent with the opportunity-cost channel: when the forgone rental income from listing a unit on Airbnb is higher, hosts set higher nightly prices to compensate. Note that when city fixed effects are included, $\overline{\text{rent}}_c$ is collinear with the city dummies; in such specifications, θ is identified from the city-level variation absorbed by the fixed effects, and we report results both with and without city fixed effects.

4.3. Model 3: City-Level Interaction Model

To characterise the bidirectional relationship between Airbnb activity and rents at the city level, we estimate a forward and a reverse aggregate regression:

$$\overline{\ln(\text{rent})}_c = a_1 + b_1 \cdot \text{airbnb\_count}_c + \mathbf{W}_c' \boldsymbol{\phi}_1 + e_{1c}, \tag{4}$$

$$\text{airbnb\_count}_c = a_2 + b_2 \cdot \overline{\ln(\text{rent})}_c + \mathbf{W}_c' \boldsymbol{\phi}_2 + e_{2c}, \tag{5}$$

where overlines denote city-level means, airbnb_count_c is the total number of Airbnb listings in city c, and $\mathbf{W}_c$ contains city-level mean bedrooms and bathrooms. The coefficients b_1 and b_2 describe the cross-city co-movement of rents and Airbnb activity. Because many of the 153 cities in the aggregated sample contribute few underlying listings, these regressions should be interpreted as descriptive rather than inferential.

4.4. Model 4: Spatial Analysis

To account for spatial dependence in rents, we estimate a spatial autoregressive (SAR) model and a spatial error model (SEM) alongside the OLS baseline. The SAR model augments the hedonic equation with a spatial lag of the dependent variable:

$$\ln(\text{rent}_i) = \alpha'' + \rho \cdot \sum_k w_{ik} \ln(\text{rent}_k) + \beta' \cdot \texttt{airbnb\_count}_{i,r} + \mathbf{X}_i' \boldsymbol{\gamma}' + \sum_c \delta_c'' \cdot \mathbf{1}[\text{city}_i = c] + \eta_i, \tag{6}$$

where $\mathbf{W} = [w_{ik}]$ is a row-standardised k-nearest-neighbour spatial weights matrix with $k = 5$, so that $\sum_k w_{ik} \ln(\text{rent}_k)$ is the mean log rent of the five nearest rental listings. The spatial autoregressive parameter ρ captures the degree to which rents co-move within a spatial neighbourhood, after controlling for observed characteristics. The SEM instead places the spatial process in the disturbance, $\eta_i = \lambda \sum_k w_{ik} \eta_k + \nu_i$, capturing spatially correlated unobservables. Both models are estimated by the feasible generalised-moments procedures of Kelejian and Prucha (1998) and Kelejian and Prucha (1999) (`GM_Lag` and `GM_Error`), which avoid the simultaneity bias that OLS estimation of Equation (6) would entail; see Anselin (1988) and LeSage and Pace (2009) for the underlying theory.

This specification serves two purposes. First, the spatial models absorb variation from spatially correlated unobservables (e.g., neighbourhood amenities that affect both rents and Airbnb desirability). Second, comparing $\hat{\beta}'$ from Equation (6) with $\hat{\beta}$ from Equation (2) provides a diagnostic for the sensitivity of the Airbnb coefficient to spatial confounders.

In addition to the spatial lag model, we examine the robustness of the Airbnb exposure measure by varying the buffer radius r across 250 m, 500 m, 1 km, and 2 km. This exercise traces out the spatial decay of the association: if the Airbnb–rent relationship is genuinely local, we expect $\hat{\beta}$ to be largest at small radii and to attenuate as the buffer expands.

4.5. Model 5: Quantile Regression

To investigate heterogeneity in the Airbnb–rent association across the conditional rent distribution, we estimate quantile regressions (Koenker and Bassett 1978; Koenker and Hallock 2001):

$$Q_\tau\left[\ln(\text{rent}_i) \mid \mathbf{X}_i, \texttt{airbnb\_count}_{i,r}\right] = \alpha_\tau + \beta_\tau \cdot \texttt{airbnb\_count}_{i,r} + \mathbf{X}_i'\boldsymbol{\gamma}_\tau + \sum_c \delta_{c,\tau} \cdot \mathbf{1}[\text{city}_i = c], \tag{7}$$

for quantile indices $\tau \in \{0.10, 0.25, 0.50, 0.75, 0.90\}$. City fixed effects are restricted to the five largest cities (with the remainder grouped) to keep the quantile estimation well conditioned. The quantile-specific coefficient β_τ measures the marginal association between Airbnb exposure and the τth quantile of log rent. If the effect of Airbnb is larger at the upper tail of the distribution, this may indicate that short-term rental activity disproportionately affects higher-quality or higher-rent segments of the market. Standard errors are the asymptotic kernel-based estimates of Koenker and Bassett (1978) as implemented in `statsmodels`.

4.6. Model 6: Machine-Learning Robustness

We complement the parametric analysis with four machine-learning methods: the LASSO (Tibshirani 1996), the elastic net (Zou and Hastie 2005), random forests (Breiman 2001), and gradient boosting (Friedman 2001), alongside an OLS benchmark; all are implemented in scikit-learn (Pedregosa et al. 2011). These models are trained to predict $\ln(\text{rent}_i)$ from the Airbnb exposure measures at the 500 m radius (count, density, mean price, entire-home share), dwelling characteristics (bedrooms, bathrooms), and geographic coordinates, and are evaluated on a held-out test set (20% of the sample) using root mean squared error (RMSE), mean absolute error (MAE), and R^2.

The machine-learning models serve three purposes. First, they provide a benchmark for the predictive accuracy of the hedonic model: if the OLS model achieves comparable R^2 to the flexible ML models, this suggests that the linear specification is not severely misspecified. Second, LASSO and elastic net coefficients reveal which variables are selected as predictors, providing a data-driven assessment of the importance of Airbnb exposure relative to dwelling characteristics. Third, random forest and gradient boosting models yield variable-importance measures (SHAP values) that quantify the contribution of each feature to predictive accuracy, without imposing functional-form assumptions.

4.7. Identification and Endogeneity

We are explicit about the identification challenges inherent in our cross-sectional design. The primary concern is that Airbnb listing density is endogenous to local rent levels and neighbourhood desirability. Three channels of endogeneity are relevant:

(i) **Reverse causality:** High rents may attract Airbnb hosts (because the opportunity cost of leaving a unit vacant is high, and short-term rental income can offset high carrying costs), creating a positive bias in the estimated $\hat{\beta}$.

(ii) **Omitted variables:** Neighbourhood amenities—such as proximity to cultural attractions, restaurants, and transit—that are unobserved in our data may simultaneously attract Airbnb guests and drive up residential rents, again biasing $\hat{\beta}$ upward.

(iii) **Spatial sorting:** Airbnb hosts may select locations based on unobserved characteristics (e.g., architectural charm, noise tolerance norms) that also affect residential rents.

Our city fixed effects control for all time-invariant, city-level confounders, including differences in local housing-market tightness, tourism infrastructure, and regulatory regimes. Within-city variation in Airbnb density is our identifying variation, and this variation is plausibly driven by both locational amenities (tourist attractiveness) and the supply of suitable housing units—neither of which is fully exogenous. We therefore interpret our estimates as conditional correlations rather than causal effects. The inclusion of spatial lags, robustness checks across buffer radii, and machine-learning diagnostics provides indirect evidence on the plausibility and stability of our estimates, but a definitive causal statement would require either panel data with suitable fixed effects or a credible instrument for Airbnb supply—neither of which is available in our setting.

4.8. Standard Errors

Throughout the analysis, we report heteroskedasticity-robust standard errors (HC1) as our baseline; Section 6 re-computes the preferred specification with standard errors clustered at the city level. For the quantile regressions, we report the asymptotic standard errors produced by the kernel-based estimator in `statsmodels`. For the machine-learning models, we report held-out test-set performance metrics, with regularisation parameters selected by 5-fold cross-validation.

5. Results

This section presents the empirical results for each of the six models described in Section 4. Throughout, we use cautious language to reflect the conditional nature of our estimates: the terms "associated with" and "conditional correlation" are used in place of "effect" or "impact," consistent with the limitations of our cross-sectional identification strategy.

5.1. Baseline Hedonic Rent Model

Table 3 reports the OLS estimates of the hedonic rent model (Equation 2). Column (1a) presents the bivariate specification with only the Airbnb count within 500 m; column (1b) adds dwelling controls (bedrooms, bathrooms, building type); and column (1c)—our preferred specification—adds city fixed effects. Columns (1d) and (1e) replace the Airbnb count with two alternative exposure measures, the Airbnb density and the share of entire-home listings within the buffer.

Table 3. Hedonic Rent Model — Baseline OLS Results

Dep. var:	**(1a)** *log_rent*	**(1b)** *log_rent*	**(1c)** *log_rent*	**(1d)** *log_rent*	**(1e)** *log_rent*
const	7.5990***	7.0390***	7.2380***	7.2380***	7.2766***
	(0.0040)	(0.0139)	(0.0677)	(0.0677)	(0.0332)
airbnb_count_500m	0.0020***	0.0048***	0.0039***		
	(0.0003)	(0.0002)	(0.0002)		
airbnb_density_500m				0.0030***	
				(0.0002)	
share_entire_home_500m					-0.0041
					(0.0107)
bedrooms		0.1118***	0.1276***	0.1276***	0.1215***
		(0.0046)	(0.0046)	(0.0046)	(0.0057)
bathrooms		0.2604***	0.2287***	0.2287***	0.3033***
		(0.0128)	(0.0126)	(0.0126)	(0.0104)
bt_House		0.0880***	0.0825***	0.0825***	0.0305
		(0.0128)	(0.0137)	(0.0137)	(0.0260)
bt_Row / Townhouse		0.1493***	0.0981***	0.0981***	0.1187*
		(0.0375)	(0.0366)	(0.0366)	(0.0695)
Observations	8303	7925	7925	7925	4370
R^2	0.0075	0.4705	0.5697	0.5697	0.4961
Adj. R^2	0.0074	0.4702	0.5554	0.5554	0.4857

Models (1c)-(1e) include city fixed effects (not shown).

Robust (HC1) standard errors in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

The coefficient on `airbnb_count_500m` is positive and statistically significant across all specifications. In the preferred specification (column 1c), each additional Airbnb listing within 500 m is associated with an approximate 0.4% increase in monthly rent, holding dwelling characteristics and city constant. This semi-elasticity is economically modest but statistically robust: at the median monthly rent of approximately $1,950, one additional nearby Airbnb listing corresponds to a rent differential of roughly $6–$10 per month. However, the practical significance compounds when one considers that many urban rental listings have 10 or more Airbnb listings within 500 m, implying cumulative differentials that are economically meaningful.

Among the control variables, dwelling size is the dominant predictor of rent: each additional bedroom is associated with approximately 11–13% higher rent, and each additional bathroom with approximately 23–30% higher rent. Building type (House and Row/Townhouse relative to Apartment) also enters significantly. City fixed effects raise the explanatory power of the model considerably, confirming substantial cross-city variation in rent levels.

The adjusted R^2 of the preferred specification is approximately 0.56, indicating that observed dwelling characteristics and city fixed effects explain a substantial share of the cross-sectional variation in rents, though a considerable residual remains, consistent with the importance of unobserved unit-specific and micro-locational factors.

5.2. Hedonic Airbnb Pricing Model

Table 4 reports the estimates of the Airbnb pricing model (Equation 3).

Contrary to the opportunity-cost hypothesis, the city-level mean rent does not enter significantly in any specification: the coefficient $\hat{\theta}$ is small and statistically indistinguishable from zero, with a negative point estimate in columns (2a) and (2b) and a positive one in column (2c). Airbnb nightly prices in our sample are therefore not systematically higher in higher-rent cities once listing characteristics are taken into account; the pricing of short-term rentals appears to be driven primarily by the properties of the listing itself rather than by conditions in the local long-term rental market.

Among the listing-level controls, dwelling size matters most: bedrooms and, especially, bathrooms are associated with significantly higher nightly prices, in line with the STR pricing literature (Wang and Nicolau 2017; Gibbs et al. 2018). Superhost status carries a significant *negative* coefficient—a notable departure from the positive reputation premia typically reported (Wang and Nicolau 2017; Ert et al. 2016)—which likely reflects composition effects: superhosts in our Quebec sample are concentrated in more modest, high-volume urban units rather than in the luxury chalet segment, and guest capacity enters negatively once size is controlled for. The entire-home indicator is positive and significant once city fixed effects are included (column 2c), consistent with whole units commanding a premium over rooms in shared dwellings within the same market (Gibbs et al. 2018).

Table 4. Hedonic Airbnb Pricing Model — OLS Results

	(2a)	**(2b)**	**(2c)**
Dep. var:	*log_price*	*log_price*	*log_price*
const	6.1958***	6.3698***	4.7948***
	(0.8696)	(0.8889)	(0.7004)
mean_rent_city	-0.0296	-0.0531	0.0307
	(0.1129)	(0.1180)	(0.1039)
bedrooms		0.0923**	0.0852*
		(0.0448)	(0.0457)
bathrooms		0.1555***	0.1226***
		(0.0428)	(0.0458)
guests_count		-0.0409**	-0.0187
		(0.0182)	(0.0190)
amenities_count		0.0008	0.0019
		(0.0022)	(0.0024)
is_superhost		-0.3467***	-0.2766***
		(0.0621)	(0.0651)
is_entire_home		-0.0722	0.1670**
		(0.0661)	(0.0749)
Observations	2437	2437	2437
R^2	0.0000	0.0300	0.1404
Adj. R^2	-0.0004	0.0272	0.0804

Model (2c) includes city fixed effects (not shown).

Robust (HC1) standard errors in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

When city fixed effects are included, the listing-level coefficients retain their signs and broad magnitudes, suggesting that the within-city pricing structure of Airbnb listings is largely independent of the cross-city rent variation.

5.3. City-Level Results

Table 5 presents the city-level regressions (Equations 4 and 5).

Table 5. City-Level Regressions

	(3-fwd)	**(3-rev)**
Dep. var:	*mean_log_rent*	*airbnb_count_city*
const	6.9341***	-40.9296
	(0.0707)	(58.2750)
airbnb_count_city	0.0001**	
	(0.0000)	
mean_bedrooms	0.0882***	-14.4332
	(0.0241)	(9.4579)
mean_bathrooms	0.3741***	11.6631*
	(0.0528)	(6.4498)
mean_log_rent		10.2254
		(10.9040)
Observations	153	153
R^2	0.4916	0.0136

Robust (HC1) standard errors in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

The forward regression shows a positive and statistically significant—though economically small—cross-city association between the total number of Airbnb listings and mean log rents: cities with more Airbnb listings tend to have somewhat higher average rents, conditional on average dwelling size. The reverse regression, by contrast, has essentially no explanatory power ($R^2 = 0.01$), and mean rent does not significantly predict city-level Airbnb counts. Although the aggregation yields 153 city-level observations, most cities contribute only a handful of underlying listings, which limits the statistical power of these regressions and amplifies the influence of individual city outliers. We interpret these results as descriptive patterns that motivate the listing-level analysis rather than as evidence of a causal relationship.

5.4. Spatial Model Results

Table 6 reports the OLS baseline alongside the spatial lag (SAR) and spatial error (SEM) models of Equation 6.

Table 6. Spatial Regression Models

	OLS	**SAR (GM_Lag)**	**SEM (GM_Error)**
Dep. var:		*log_rent*	
const	7.1114***	6.0767***	7.0905***
	(0.0211)	(0.1514)	(0.0328)
airbnb_count_500m	0.0038***	0.0034***	0.0038***
	(0.0002)	(0.0002)	(0.0004)
bedrooms	0.1194***	0.1220***	0.1546***
	(0.0045)	(0.0032)	(0.0030)
bathrooms	0.2432***	0.2285***	0.1795***
	(0.0123)	(0.0063)	(0.0057)
bt_House	0.1342***	0.1188***	0.1299***
	(0.0131)	(0.0101)	(0.0097)
bt_Row / Townhouse	0.1782***	0.1604***	0.1447***
	(0.0359)	(0.0313)	(0.0296)
W_log_rent		0.1367***	
		(0.0199)	
λ (spatial error)			0.5393
Observations	7925	7925	7925
R^2 / pseudo-R^2	0.5081	0.5503	0.4972

Standard errors in parentheses. SAR estimated via GM_Lag; SEM via GM_Error. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$. City FE included but not shown.

The spatial autoregressive coefficient ($\hat{\rho}$, the coefficient on $W \cdot \ln(\text{rent})$) is positive and strongly significant, and the spatial error parameter $\hat{\lambda}$ in the SEM is large (0.54), confirming substantial spatial dependence in rents: the rents of nearby listings carry systematic information about a given listing's rent, even after controlling for dwelling characteristics and city fixed effects. This finding is consistent with the large literature on spatial autocorrelation in housing markets.

Importantly, the coefficient on `airbnb_count_500m` ($\hat{\beta}'$) remains positive and statistically significant in both spatial specifications: it is mildly attenuated in the SAR model (0.0034 versus 0.0038 in the OLS baseline) and essentially unchanged in the SEM. The mild attenuation is expected: to the extent that the Airbnb coefficient in the baseline model partly reflected spatially correlated omitted variables (captured by the spatial terms), the spatial models provide more conservative estimates. The persistence of a significant positive association strengthens confidence that the Airbnb–rent correlation is not solely an artefact of spatial confounding.

Figure 4 illustrates the spatial decay of the Airbnb coefficient across buffer radii.

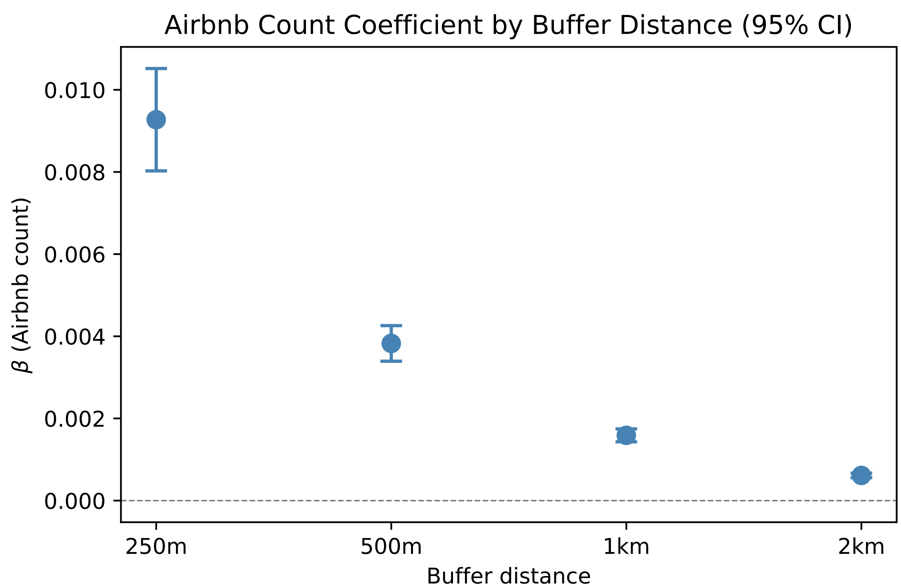


Figure 4. Estimated Airbnb Coefficient by Buffer Radius

The per-listing coefficient is largest at the 250 m radius and declines monotonically as the buffer expands, consistent with a localised association between Airbnb activity and rents. At the 2 km radius, the coefficient remains positive but is substantially smaller in magnitude, reflecting the dilution of the Airbnb signal over a larger geographic area. This spatial decay pattern is consistent with prior findings in the literature and suggests that the Airbnb–rent association operates at a highly localised scale.

5.5. Quantile Regression Results

Table 7 and Figure 5 present the quantile regression estimates of β_τ for $\tau \in \{0.10, 0.25, 0.50, 0.75, 0.90\}$.

Table 7. Quantile Regression Results — Coefficient on Airbnb Count (500m)

	$\tau = 0.10$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.90$	**OLS**
Dep. var:	*log_rent*	*log_rent*	*log_rent*	*log_rent*	*log_rent*	*log_rent*
const	6.8249***	6.9035***	7.0339***	7.1237***	7.2461***	7.0508***
	(0.0258)	(0.0211)	(0.0209)	(0.0221)	(0.0328)	(0.0197)
airbnb_count_500m	0.0037***	0.0034***	0.0035***	0.0041***	0.0047***	0.0038***
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0003)	(0.0002)
bedrooms	0.1236***	0.1034***	0.1071***	0.1360***	0.1418***	0.1152***
	(0.0058)	(0.0043)	(0.0039)	(0.0038)	(0.0055)	(0.0046)
bathrooms	0.2411***	0.2796***	0.2748***	0.2669***	0.2807***	0.2571***
	(0.0103)	(0.0076)	(0.0072)	(0.0076)	(0.0115)	(0.0125)
bt_House	0.0608***	0.1079***	0.1382***	0.1079***	0.1094***	0.1201***
	(0.0139)	(0.0119)	(0.0120)	(0.0131)	(0.0205)	(0.0131)
bt_Row / Townhouse	0.1310***	0.0858**	0.1602***	0.2501***	0.2274***	0.1753***
	(0.0468)	(0.0383)	(0.0383)	(0.0417)	(0.0625)	(0.0378)
City FE (top 5)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7925	7925	7925	7925	7925	7925
(Pseudo-)R^2	0.2251	0.2529	0.2976	0.3294	0.3402	0.4849

Standard errors in parentheses. OLS uses HC1 robust SE. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$. City FE limited to the 5 largest cities (others grouped).

The quantile regression results reveal meaningful heterogeneity in the Airbnb–rent association across the conditional rent distribution. The coefficient is positive and statistically significant at every quantile considered. It is roughly flat over the lower half of the distribution (0.0037 at $\tau = 0.10$, 0.0034–0.0035 at $\tau = 0.25$ and $\tau = 0.50$) and then rises in the upper tail, reaching 0.0041 at $\tau = 0.75$ and its largest value, 0.0047, at $\tau = 0.90$—roughly 25% above the OLS estimate.

This pattern is economically interpretable. High-rent listings—typically located in desirable neighbourhoods with strong tourist appeal—are precisely the locations where Airbnb activity is most intense and where the supply-withdrawal mechanism is most plausible. The finding that implicit prices vary across the conditional distribution echoes the broader hedonic evidence (Zietz et al. 2008; McMillen 2008), and extends it to STR exposure: the association between Airbnb and rents, while present throughout the distribution, is strongest in the upper segment of the rental market.

5.6. Machine-Learning Robustness

Table 8 reports the out-of-sample predictive performance of the four machine-learning models alongside the OLS benchmark.

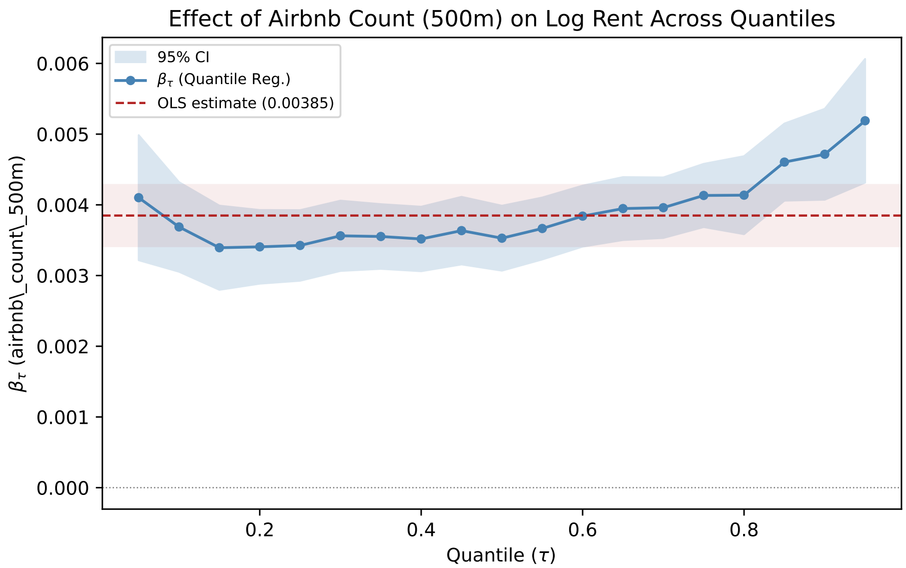


Figure 5. Quantile Regression Coefficients on Airbnb Count (500m) with 95% Confidence Intervals

Table 8. Machine-Learning Model Performance (Test Set)

Model	R^2 (Train)	R^2 (Test)	RMSE (Test)	MAE (Test)
OLS	0.4741	0.5094	0.2242	0.1727
LASSO	0.4740	0.5099	0.2241	0.1725
Elastic Net	0.4740	0.5099	0.2241	0.1725
Random Forest	0.9092	0.7079	0.1730	0.1284
Gradient Boosting	0.8762	0.6935	0.1772	0.1319

Notes: OLS, LASSO, and Elastic Net use standardized features. Tree-based models use raw features. Train/test split is 80/20 with random_state=42.

The random forest achieves the highest R^2 on the test set (0.71), closely followed by gradient boosting (0.69); the regularised linear models (LASSO and elastic net) perform essentially on par with OLS (test $R^2 \approx 0.51$). The gap between the tree-based and linear models indicates that nonlinear relationships and interactions among covariates—most plausibly involving the geographic coordinates—contribute meaningfully to rent variation beyond what the linear hedonic model captures. The linear specification nevertheless accounts for the bulk of the explainable variation and remains a reasonable approximation for inference on the Airbnb exposure coefficient.

Figure 6 displays the mean absolute SHAP values from the gradient boosting model.

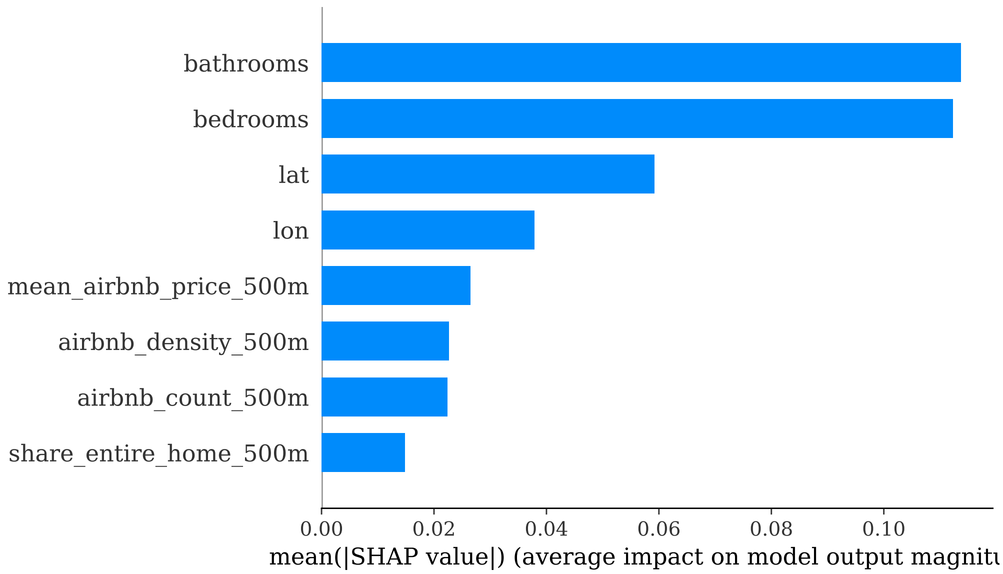


Figure 6. Feature Importance from Gradient Boosting Model

Dwelling characteristics (bathrooms, bedrooms) and geographic coordinates are consistently the most important predictors of rent. The Airbnb exposure variables—mean nearby price, density, count, and entire-home share—each rank as moderately important features and jointly account for a non-trivial share of predictive power, confirming that Airbnb exposure carries information beyond what is captured by dwelling size and raw location. The LASSO and elastic net models retain the Airbnb count variable with a positive coefficient, consistent with the OLS results.

Figure 7 presents a SHAP summary plot, providing a more granular view of how each feature contributes to individual predictions.

The SHAP analysis confirms that higher Airbnb counts are associated with positive contributions to predicted rent, consistent with the parametric results. The distribution of SHAP values for the Airbnb count variable shows a rightward shift for listings with high Airbnb exposure, reinforcing the finding that nearby short-term rental activity is associated with higher rents.

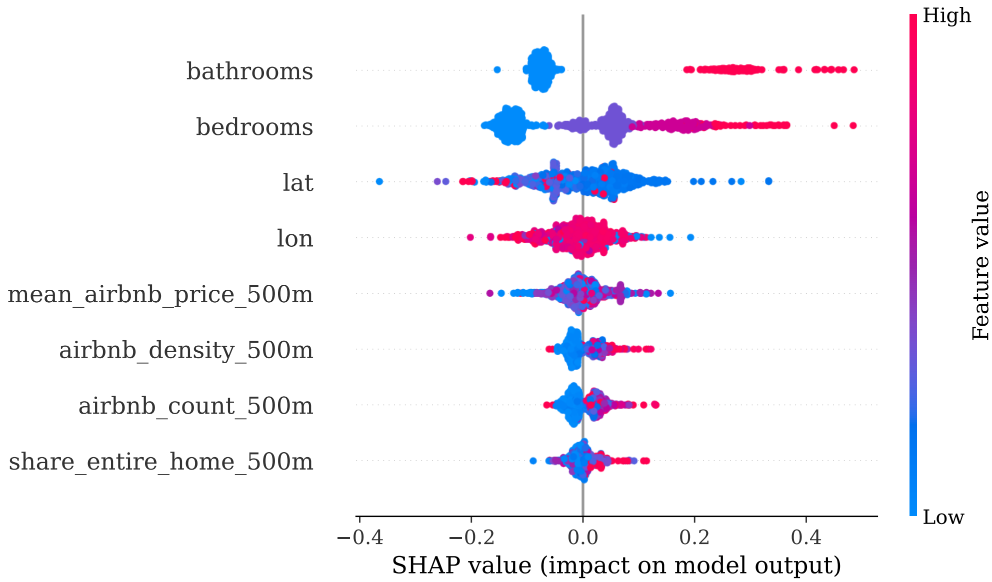


Figure 7. SHAP Summary Plot from Gradient Boosting Model

6. Robustness Checks

This section presents a battery of robustness tests designed to assess the sensitivity of our baseline estimates to alternative specifications, sample definitions, and exposure measures.

6.1. Buffer Radius Robustness

Our baseline specification uses a 500 m buffer to define Airbnb exposure. Table 9 reports the coefficient on the Airbnb count variable for buffer radii of 250 m, 500 m, 1 km, and 2 km, estimated using the same hedonic specification (Equation 2) with the full set of dwelling controls and city fixed effects.

Table 9. Buffer Radius Robustness — Hedonic Rent Model

	250m	**500m**	**1km**	**2km**
Dep. var:	*log_rent*	*log_rent*	*log_rent*	*log_rent*
Airbnb count	0.009272***	0.003826***	0.001589***	0.000612***
	(0.000636)	(0.000221)	(0.000080)	(0.000028)
Controls	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Observations	7,925	7,925	7,925	7,925
R^2	0.5026	0.5081	0.5167	0.5205

Robust (HC1) standard errors in parentheses. Controls: bedrooms, bathrooms, building-type dummies. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

The results confirm the spatial decay pattern illustrated in Figure 4. The per-listing coefficient is largest at the 250 m radius, where the spatial signal is most concentrated, and decreases monotonically as the buffer expands—mechanically so, since a listing counted within a wide buffer is a weaker proxy for immediate proximity than one counted within a narrow buffer. At all radii, the coefficient is positive and statistically significant at the 1% level. The consistency of the positive association across radii provides reassurance that the finding is not an artefact of the particular buffer choice.

Figure 8 visualises the coefficient estimates and their 95% confidence intervals across buffer radii.

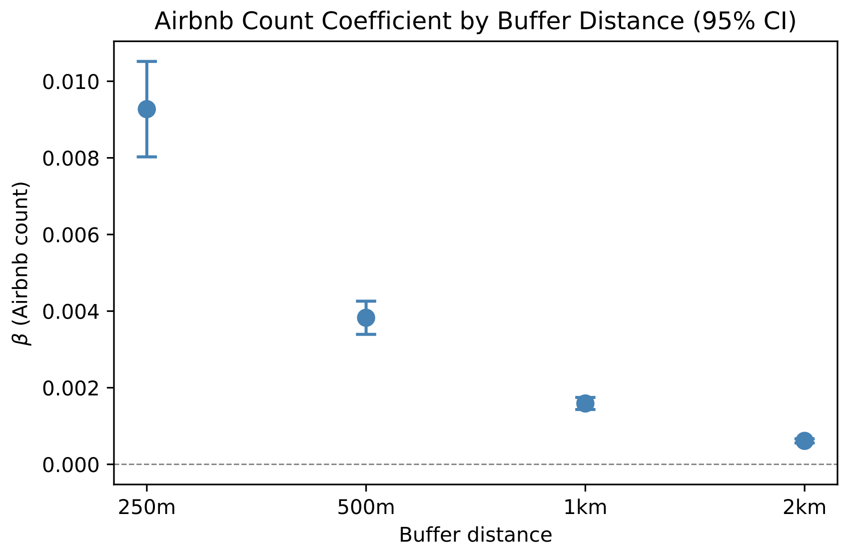


Figure 8. Airbnb Count Coefficient Across Buffer Radii (with 95% CI)

6.2. Alternative Exposure Measures

Our baseline uses the raw count of Airbnb listings as the exposure measure. We consider two alternative measures to assess whether the results are sensitive to the functional form of the exposure variable:

(i) **Airbnb density:** the count divided by the buffer area (πr^2 in km^2), which normalises for the geometric expansion of the buffer.

(ii) **Share of entire-home listings:** the fraction of Airbnb listings within the buffer that are classified as entire-home properties (House, Cabin/Chalet, Condo), capturing the composition of short-term rental activity.

These alternative specifications are reported as Models (1d) and (1e) in Table 3.

The Airbnb density measure yields qualitatively identical results to the raw count, with a positive and significant coefficient—as expected, since at a fixed radius the density is a rescaling of the count. The share of entire-home listings, by contrast, is *not* significantly associated with rents (the point estimate is small and negative): conditional on dwelling characteristics and city fixed effects, it is the intensity of nearby Airbnb activity, rather than its compositional tilt toward entire homes, that co-varies with rents in our data. This finding does not support the compositional prediction of the supply-withdrawal hypothesis, under which entire-home listings—the closest substitutes for long-term rental units—should matter most; we return to this point in Section 7.

6.3. Subsample Analysis

To assess whether the baseline results are driven by a particular geographic segment or building type, we estimate the hedonic rent model separately for the following subsamples:

(i) **Montreal vs. non-Montreal:** Montreal dominates both datasets and is the primary tourist destination. The Airbnb–rent association may be stronger in Montreal, where short-term rental activity is most concentrated, or weaker if the city's larger and more liquid rental market is better able to absorb the supply shock.

(ii) **Apartments vs. Houses:** Apartments constitute the vast majority of rental listings and are arguably the closest substitute for Airbnb rental units. Houses, by contrast, are a more heterogeneous category and may operate in a partially segmented market.

Table 10 presents the results.

Table 10. Subsample Analysis — Hedonic Rent Model

	(1) Montreal	(2) Outside Mtl	(3) Apartments	(4) Houses
Airbnb count (500m)	0.003865***	0.003046*	0.003860***	0.006699***
	(0.000226)	(0.001580)	(0.000226)	(0.001995)
N	4,578	3,347	7,136	789
R^2	0.4587	0.6959	0.5046	0.6403
City FE	Yes	Yes	Yes	Yes

Notes: Robust standard errors (HC1) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All specifications include city fixed effects, bedrooms, and bathrooms as controls.

The Montreal subsample yields a positive and significant Airbnb coefficient that is similar in magnitude to the full-sample estimate, confirming that the baseline results are not driven solely by the inclusion of non-Montreal observations. The non-Montreal subsample also produces a positive coefficient, though with lower precision due to the smaller sample size and the greater heterogeneity of non-Montreal markets (which include both urban centres and resort communities).

The apartment subsample closely mirrors the full-sample results, which is unsurprising given that apartments constitute approximately 90% of the rental sample. The house subsample yields a coefficient that is larger than the full-sample estimate (0.0067) and statistically significant despite the much smaller sample, though its wider confidence interval reflects the greater heterogeneity of house rentals.

6.4. Outlier Sensitivity

To assess the sensitivity of the results to extreme values, we re-estimate the baseline model on two trimmed samples: one that retains only listings with monthly rents between the 5th and 95th percentiles, and one that drops the top and bottom 1% of observations by Airbnb count within 500 m (i.e., listings in the most Airbnb-saturated locations).

Table 11. Outlier Sensitivity — Hedonic Rent Model

	(1) Full Sample	(2) Rent 5/95	(3) Airbnb 1/99
Airbnb count (500m)	0.003891***	0.003057***	0.004222***
	(0.000225)	(0.000186)	(0.000242)
N	7,925	7,254	7,868
R^2	0.5664	0.4966	0.5673
City FE	Yes	Yes	Yes

Notes: Robust standard errors (HC1) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column (2) winsorizes monthly rent at the 5th and 95th percentiles. Column (3) trims the top and bottom 1% of Airbnb listing counts.

The results are stable across the outlier-handling approaches. Trimming the rent distribution at the 5th and 95th percentiles reduces the point estimate by roughly a fifth (from 0.0039 to 0.0031)—consistent with the quantile-regression finding that the association is strongest in the tails—while trimming extreme Airbnb counts slightly increases it (0.0042). In every case the coefficient remains positive and significant at the 1% level, providing confidence that the baseline results are not driven by a small number of influential observations.

6.5. Additional Specification Checks

Table 12 subjects the preferred specification to four further checks. Column (R1) reproduces the baseline for reference.

Clustered standard errors. Column (R2) re-computes the baseline standard errors clustering by city (153 clusters). HC1 standard errors assume independence across observations, which is questionable when listings are spatially concentrated. In our data, city-level clustering in fact *tightens* the standard error on the Airbnb coefficient (0.0001 versus 0.0002 under HC1), so the baseline significance levels are, if anything, conservative on this dimension. Clustering does widen the confidence interval on some controls (bathrooms), as expected.

Table 12. Extended Robustness — Preferred Hedonic Specification

Dep. var:	**(R1)** *log_rent*	**(R2)** *log_rent*	**(R3)** *log_rent*	**(R4)** *log_rent*	**(R5)** *log_rent*
Airbnb count (500m)	0.0039***	0.0039***	0.0032***		0.0013***
	(0.0002)	(0.0001)	(0.0003)		(0.0003)
log(1 + Airbnb count 500m)				0.0484***	
				(0.0025)	
Airbnb count (500m–1km ring)					0.0018***
					(0.0002)
Bedrooms	0.1276***	0.1276***	0.1271***	0.1296***	0.1322***
	(0.0046)	(0.0043)	(0.0076)	(0.0046)	(0.0047)
Bathrooms	0.2287***	0.2287***	0.2103***	0.2274***	0.2246***
	(0.0126)	(0.0291)	(0.0202)	(0.0125)	(0.0125)
Interior size (100 sq ft)			0.0044***		
			(0.0017)		
Standard errors	HC1	City cluster	HC1	HC1	HC1
Building type & city FE	Yes	Yes	Yes	Yes	Yes
Observations	7,925	7,925	3,531	7,925	7,925
R^2	0.5697	0.5697	0.6044	0.5726	0.5786

Standard errors in parentheses: HC1 except column (R2), which clusters by city. All columns control for bedrooms, bathrooms, building type, and city fixed effects; column (R3) adds interior size (per 100 sq ft) on the subsample reporting it. Column (R5) includes the 500m count and the count in the 500m–1km annulus jointly. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Interior size. Column (R3) adds the interior floor area (per 100 sq ft) as a control on the subsample of 3,531 listings that report it. Size enters positively and significantly, and the Airbnb coefficient remains positive and highly significant at 0.0032—about a fifth smaller than the baseline, consistent with size absorbing some quality variation, but leaving the qualitative conclusion intact.

Functional form. Column (R4) replaces the linear count with $\ln(1 + \texttt{airbnb\_count\_500m})$. The elasticity-style coefficient (0.048) is positive and highly significant: doubling the nearby Airbnb count is associated with roughly 3.4% higher rent. The fit improves marginally over the linear form, indicating mild concavity in the exposure–rent relationship.

Ring exposure. Column (R5) includes the 500 m count jointly with the count in the 500 m–1 km annulus. Both enter positively and significantly. Splitting the exposure this way attenuates the inner-buffer coefficient (0.0013), and the annulus coefficient (0.0018) is of comparable magnitude per listing: the association is therefore not confined to the immediate 500 m but extends to the kilometre scale, consistent with the buffer-radius results above, where the total (not per-listing) association accumulates over wider areas. Because the two counts are strongly correlated, the individual coefficients in this specification should be read as a decomposition of a common neighbourhood signal rather than as separate causal gradients.

6.6. Leave-One-City-Out Sensitivity

Figure 9 re-estimates the preferred specification excluding, in turn, each of the eight cities with the most rental listings.

The coefficient is remarkably stable, ranging from 0.0038 to 0.0040 across exclusions. Notably, excluding Montreal—which accounts for more than half of the estimation sample—leaves the point estimate essentially unchanged (0.0038), albeit with a wider confidence interval; the association is thus not an artefact of any single market.

6.7. Summary of Robustness

Figure 10 presents a coefficient plot summarising the Airbnb count coefficient across all robustness specifications.

Across all specifications—varying the buffer radius, the exposure measure, the functional form, the control set, the sample, the outlier treatment, and the standard-error computation—the estimated association between Airbnb presence and residential rents is consistently positive. While the magnitude varies across specifications, the qualitative conclusion is robust: higher Airbnb density in the immediate vicinity of a rental listing is associated with higher monthly rents, conditional on observable dwelling characteristics and city-level heterogeneity.

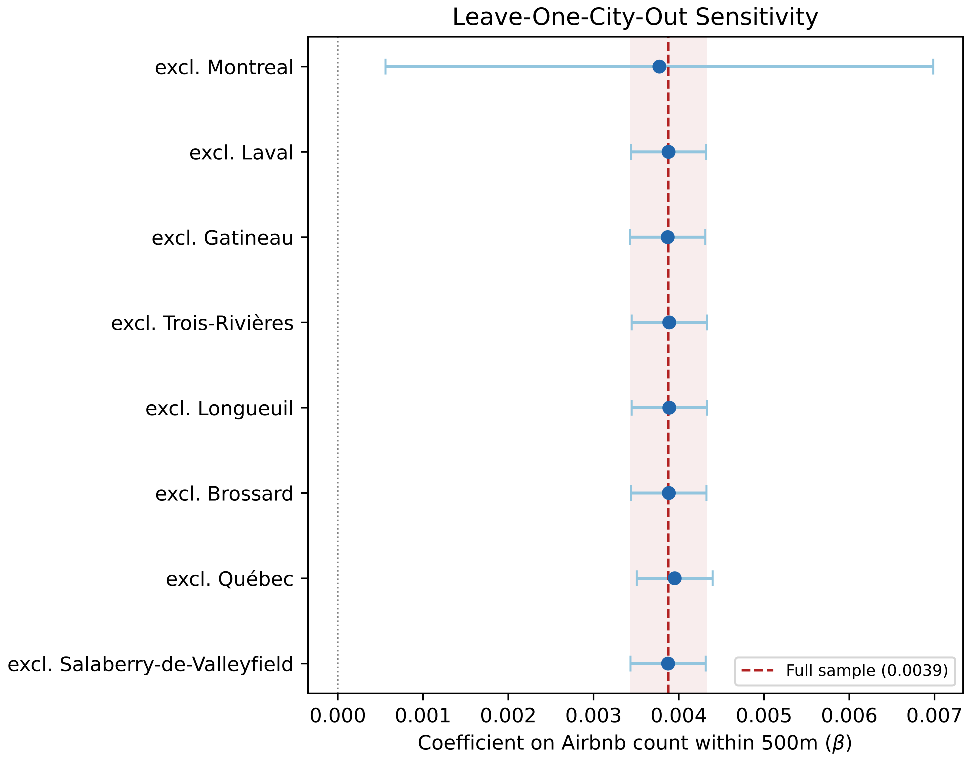


Figure 9. Leave-One-City-Out Sensitivity of the Airbnb Coefficient

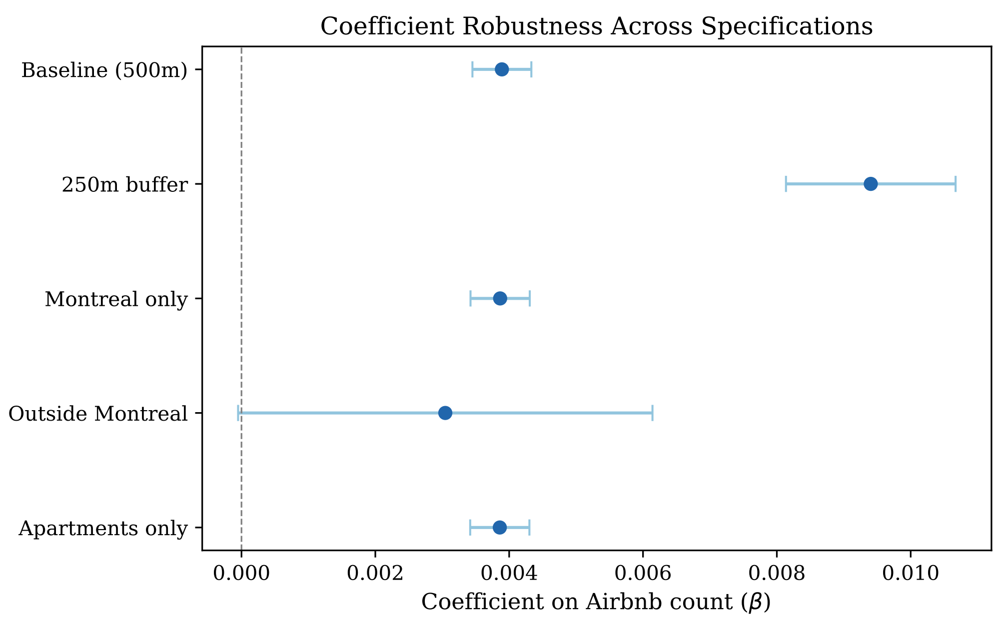


Figure 10. Summary of Airbnb Count Coefficients Across Specifications

7. Discussion

This section places the empirical findings in the context of the international literature, interprets them for housing policy, discusses the mechanisms that may underlie the observed associations, and addresses the limitations of the analysis.

7.1. Findings in the Context of the Literature

How do our magnitudes compare with prior work? Three benchmarks are instructive. First, Duso et al. (2024), using regulatory shocks in Berlin, estimate that each nearby Airbnb listing raises asking rents by roughly 7–13 euros per month; our preferred estimate implies \$6–10 CAD per listing at the median Quebec rent—the same order of magnitude, obtained from an entirely different market and design. Second, Koster et al. (2021) find that Los Angeles ordinances which halved Airbnb activity lowered local rents by about 2%; read through our semi-elasticity, removing half of the mean nearby exposure in our sample would predict a rent differential of a comparable magnitude in Airbnb-dense neighbourhoods. Third, Horn and Merante (2017) report that a one-standard-deviation increase in Airbnb density in Boston is associated with 0.4% higher asking rents; our per-listing coefficient of 0.4%, combined with the dispersion of exposure in our data, implies a somewhat larger standardised association, consistent with Quebec's rental stock being more concentrated in dense, Airbnb-exposed neighbourhoods. Our estimates also sit comfortably within the range documented across European markets (Garcia-López et al. 2020; Franco and Santos 2021; Ayouba et al. 2020).

Two disagreements with the literature deserve emphasis rather than concealment. First, unlike most STR pricing studies (Wang and Nicolau 2017; Ert et al. 2016; Gibbs et al. 2018), we find a superhost *discount* and no significant association between city-level rents and nightly prices (Model 2); we attribute both to composition—Quebec's STR market mixes urban apartments with resort chalets, and the cross-city dimension is dominated by resort communities where nightly prices are high but long-term rents are not. Second, the compositional prediction of the supply-withdrawal hypothesis—that entire-home listings should matter most (Horn and Merante 2017; Combs et al. 2020)—is not supported in our data: the entire-home share within the buffer is statistically indistinguishable from zero once fixed effects are included (Model 1e). Either our property-type proxy for entire-home status is too coarse, or in Quebec the intensity of nearby STR activity matters more than its composition. Both null results are informative for the Canadian debate precisely because they were not guaranteed by the design.

7.2. Policy Implications for Housing Affordability

Our finding of a positive association between Airbnb presence and residential rents, while not causal in the strict econometric sense, carries implications for housing affordability policy. If the association reflects, even in part, a genuine supply-withdrawal mechanism—whereby landlords convert long-term rental units into short-term Airbnb listings—then the cumulative effect on rents in high-tourism neighbourhoods could be substantial. At our estimated semi-elasticity of 0.3–0.5% per additional Airbnb listing within 500 m, a neighbourhood with 20 nearby Airbnb listings would be associated with 6–10% higher rents relative to an otherwise identical dwelling in an Airbnb-free area, all else equal.

For policymakers in Quebec and Montreal, these estimates suggest that short-term rental activity is a factor—though certainly not the only or dominant factor—in the rental affordability equation. The magnitude of the association is modest relative to the contribution of dwelling characteristics (bedrooms, bathrooms) and location (city fixed effects), but it is non-trivial in a market where vacancy rates are historically low and modest rent increases impose real burdens on lower-income tenants.

7.3. Short-Term Rental Regulation

The province of Quebec requires short-term rental operators to register with the Corporation de l'industrie touristique du Québec (CITQ) and imposes minimum standards on registered operators. The City of Montreal has supplemented these provincial requirements with additional restrictions, including limits on the number of nights per year that a primary residence can be rented on a short-term basis and outright prohibitions on non-owner-occupied short-term rentals in certain boroughs.

Our results, while descriptive, provide empirical grounding for the policy rationale underlying these regulations. The spatial decay of the Airbnb–rent association (Section 5) suggests that the effects are highly localised, which supports geographically targeted interventions (e.g., borough-level restrictions) rather than blanket province-wide regulations. The quantile regression results further suggest that regulatory attention might focus on high-rent neighbourhoods, where the association is strongest and where the risk of supply withdrawal is most acute.

However, regulation involves tradeoffs. Short-term rentals generate income for hosts, tax revenue for municipalities, and consumer surplus for travellers. Overly restrictive regulation may push short-term rental activity underground, reduce tourism revenues, and impose compliance costs on casual hosts who rent their primary residence occasionally. A well-calibrated regulatory framework would balance these considerations by targeting commercial operators—those who manage multiple entire-home listings—while preserving the ability of residents to engage in occasional home-sharing.

7.4. Professional Hosts and Commercialisation

The literature has documented the increasing commercialisation of Airbnb, with a growing share of listings operated by professional, multi-listing hosts (Ke 2017; Wachsmuth and Weisler 2018). Unfortunately, our data do not contain host identifiers, precluding direct measurement of multi-listing activity. The compositional evidence we can bring to bear is, moreover, not supportive of a simple commercialisation channel: the share of entire-home listings within the buffer—the listings most likely to be operated commercially—is not significantly associated with rents once dwelling characteristics and city fixed effects are controlled for (Model 1e in Table 3). In our cross-section, it is the overall intensity of nearby Airbnb activity, rather than its compositional tilt toward entire homes, that co-varies with rents. One interpretation is that our property-type proxy for entire-home status is too coarse to isolate commercial operations; another is that, within Quebec's market, casual and commercial listings are sufficiently co-located that composition adds little signal beyond the count.

Future research with host-level data could decompose the Airbnb–rent association into contributions from commercial versus casual hosts, providing a sharper evidence base for regulatory targeting.

7.5. The Tourism–Housing Tradeoff

At a broader level, the Airbnb–rent relationship exemplifies a fundamental tension in urban policy: the tradeoff between tourism-driven economic activity and residential affordability. Cities like Montreal derive substantial economic benefits from tourism—employment in hospitality, retail, and cultural sectors; tax revenues; and global visibility—but these benefits are unevenly distributed, while the costs of tourism-driven housing pressure fall disproportionately on renters. Short-term rental platforms amplify this tension by enabling the conversion of residential housing into tourism infrastructure at the level of individual units, bypassing the traditional regulatory apparatus that governs hotel and commercial accommodation.

Our findings suggest that this tension is empirically present in the Quebec context, though its magnitude is moderate. The policy challenge is to design regulatory frameworks that capture the benefits of short-term rental activity while mitigating its externalities on the residential rental market—a challenge that requires ongoing empirical monitoring as the platform economy continues to evolve.

7.6. Limitations

We reiterate and expand upon the key limitations of our analysis:

(i) **Cross-sectional identification:** Our data represent a single point in time. We cannot distinguish the causal effect of Airbnb on rents from reverse causality (high rents attracting Airbnb hosts) or confounding by unobserved neighbourhood characteristics. A credible causal estimate would require panel data—ideally combined with a policy shock that exogenously shifted Airbnb supply in some locations but not others, as in Koster et al. (2021), Valentin (2021), and Duso et al. (2024)—or a valid instrumental variable for Airbnb penetration in the spirit of Barron et al. (2021). That said, the consistency of our magnitudes with these designs (Section 7.1) suggests the cross-sectional bias, whatever its direction, is unlikely to be dramatic.

(ii) **No host-level data:** The absence of host identifiers prevents us from identifying multi-listing operators and from distinguishing professional from casual hosts, a distinction that is central to the commercialisation debate (Ke 2017; Combs et al. 2020) and to well-targeted regulation (Gurran and Phibbs 2017).

(iii) **Asking vs. transacted rents:** Our dependent variable measures asking rents on Realtor.ca rather than actual contract rents. If asking rents systematically overstate transacted rents (due to landlord bargaining power or strategic posting), our estimates may be biased, though the direction of this bias is ambiguous. We note that several benchmark studies share this feature (Horn and Merante 2017; Duso et al. 2024), which aids comparability even if it does not remove the concern.

(iv) **Exposure measurement error and platform coverage:** Both datasets are scraped from specific platforms and may not capture the universe of short-term or long-term rental listings. Alternative STR platforms (e.g., VRBO, Booking.com) are not included, scraped snapshots miss delisted or calendar-blocked units, and the Realtor.ca data may underrepresent informal or unposted rentals. To the extent that measurement error in the buffer counts is classical, it attenuates $\hat{\beta}$ toward zero, making our estimates conservative; non-classical error correlated with neighbourhood type cannot be ruled out.

(v) **Buffer choice and aggregation:** Any distance-based exposure measure faces a modifiable-areal-unit-style sensitivity: results could, in principle, depend on the radii chosen. We address this directly by reporting four radii, a ring decomposition, density and logarithmic functional forms, and continuous-decay-consistent patterns (Section 6), but the discreteness of any buffer set remains a simplification of the underlying spatial process (Can 1992).

(vi) **External validity:** Our sample is dominated by Montreal and a handful of other Quebec cities and resort areas. The findings may not generalise to other Canadian provinces or to rural housing markets where short-term rental dynamics differ—a caution reinforced by the cross-market heterogeneity documented in France (Ayouba et al. 2020) and Canada (Combs et al. 2020).

(vii) **Static analysis:** Our cross-sectional framework cannot capture dynamic adjustments in the housing market, such as the construction of new supply in response to rising rents (Glaeser et al. 2006; Saiz 2010) or the exit of Airbnb hosts in response to regulatory pressure.

These limitations motivate the cautious interpretive stance adopted throughout the paper and underscore the need for future research with richer data structures.

7.7. Future Work

Four extensions strike us as most valuable, in ascending order of data requirements. First, replicating the buffer-based design on repeated scrapes of the same platforms would deliver a listing-level panel for Quebec, enabling fixed-effects and event-study designs around the province's evolving registration regime. Second, Quebec's 2023 tightening of STR registration enforcement offers a natural policy discontinuity analogous to those exploited in Los Angeles (Koster et al. 2021), New Orleans (Valentin 2021), and Berlin (Duso et al. 2024); combining it with our exposure measures is the most direct route to causal estimates for Canada. Third, host-resolved data would permit separating commercial from casual supply, testing the commercialisation channel that our entire-home-share proxy cannot (Ke 2017; Combs et al. 2020). Fourth, integrating administrative contract rents (e.g., lease registry data) would resolve the asking-rent measurement concern and allow distributional welfare analysis in the spirit of Li et al. (2022).

8. Conclusion

This paper has investigated the relationship between Airbnb short-term rental activity and residential rents in Quebec, Canada, using cross-sectional microdata comprising approximately 5,000 Airbnb listings and 8,300 residential rental listings. Employing a hedonic pricing framework augmented with spatial analysis, quantile regressions, and machine-learning benchmarks, we have documented a consistent positive association between nearby Airbnb presence and monthly rents.

Our key findings are as follows. First, each additional Airbnb listing within 500 metres of a rental unit is associated with an approximate 0.4% increase in monthly rent, conditional on dwelling characteristics and city fixed effects—an estimate that is stable across spatial and trimmed-sample specifications (roughly 0.3–0.5%). Second, this association exhibits spatial decay, with the per-listing coefficient largest at narrow buffer radii and attenuating at wider distances. Third, quantile regressions reveal that the association, while significant throughout the distribution, is strongest at the upper tail, suggesting that high-rent segments of the market are most strongly linked to Airbnb activity. Fourth, a complementary hedonic model of Airbnb nightly prices indicates that short-term rental pricing is driven primarily by listing characteristics; city-level mean rents carry no significant premium, offering no support for a simple opportunity-cost pricing channel in our cross-section. Fifth, machine-learning models confirm the predictive relevance of Airbnb exposure variables and the adequacy of the linear hedonic specification for inference.

The principal contribution of this paper is to provide the first granular, listing-level econometric analysis of the Airbnb–rent nexus in Quebec—extending a Canadian evidence base that was previously descriptive (Combs et al. 2020; Grisdale 2021)—leveraging precise geographic coordinates for exact distance-based spatial matching. By applying multiple econometric and machine-learning methods to a single dataset, we demonstrate the consistency of the finding across methodological frameworks and provide a rich set of robustness checks that characterise the sensitivity of the estimates to alternative specifications. Notably, our per-listing magnitude is of the same order as those recovered from regulatory quasi-experiments in Berlin and Los Angeles (Duso et al. 2024; Koster et al. 2021), which we read as mutual corroboration between designs.

We are candid about the limitations of our analysis. The cross-sectional nature of the data precludes causal identification: the positive association between Airbnb density and rents may reflect reverse causality, omitted neighbourhood characteristics, or spatial sorting, rather than—or in addition to—a genuine supply-withdrawal effect. Establishing causality in this domain requires panel data combined with plausibly exogenous variation in Airbnb supply, such as a regulatory discontinuity or a natural experiment. Our results should therefore be interpreted as well-controlled conditional correlations that are consistent with the supply-withdrawal hypothesis but do not definitively confirm it.

Several directions for future research emerge from this analysis. First, the construction of panel data—tracking the entry and exit of Airbnb listings and the evolution of rents over time at the neighbourhood level—would enable difference-in-differences or event-study designs that can more credibly isolate the causal effect. Second, the exploitation of regulatory shocks—such as the tightening of Montreal's short-term rental regulations or the introduction of provincial registration requirements—would provide natural-experiment variation for causal inference. Third, the integration of host-level data would allow researchers to distinguish the effects of commercial multi-listing operators from those of casual home-sharers, sharpening the policy relevance of the analysis. Fourth, extending the geographic scope to include other Canadian cities would improve external validity and enable cross-city comparisons of regulatory effectiveness.

In sum, our findings add to a growing body of evidence—now spanning correlational, quasi-experimental, and structural designs (Barron et al. 2021; Koster et al. 2021; Li et al. 2022)—that short-term rental platforms are associated with higher residential rents, at least in localities with significant tourist appeal. While the estimated magnitudes are modest at the per-listing level, their cumulative significance in high-tourism neighbourhoods—combined with the documented concentration of effects at the upper end of the rent distribution—underscores the importance of evidence-based regulatory frameworks that balance the economic benefits of the platform economy with the imperative of housing affordability.

References

Agyeman, J., McLaren, D., and Schaefer-Borrego, A. (2020). Airbnb and the undoing of mid-twentieth century housing policy. *Urban Studies*, 57(14):2998–3015.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht.

Athey, S. and Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11:685–725.

Ayouba, K., Breuillé, M.-L., Grivault, C., and Le Gallo, J. (2020). Does Airbnb disrupt the private rental market? An empirical analysis for French cities. *International Regional Science Review*, 43(1–2):76–104.

Barron, K., Kung, E., and Proserpio, D. (2021). The effect of home-sharing on house prices and rents: Evidence from Airbnb. *Marketing Science*, 40(1):23–47.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, 22(3):453–474.

Cocola-Gant, A. (2016). Holiday rentals: The new gentrification battlefront. *Sociological Research Online*, 21(3):112–120.

Combs, J., Kerrigan, D., and Wachsmuth, D. (2020). Short-term rentals in Canada: Uneven growth, uneven impacts. *Canadian Journal of Urban Research*, 29(1):119–135.

Dubin, R. A. (1988). Estimation of regression coefficients in the presence of spatially autocorrelated error terms. *Review of Economics and Statistics*, 70(3):466–474.

Duso, T., Michelsen, C., Schäfer, M., and Tran, K. D. (2024). Airbnb and rental markets: Evidence from Berlin. *Regional Science and Urban Economics*, 106:104007.

Einav, L., Farronato, C., and Levin, J. (2016). Peer-to-peer markets. *Annual Review of Economics*, 8:615–635.

Ert, E., Fleischer, A., and Magen, N. (2016). Trust and reputation in the sharing economy: The role of personal photos in Airbnb. *Tourism Management*, 55:62–73.

Farronato, C. and Fradkin, A. (2022). The welfare effects of peer entry: The case of Airbnb and the accommodation industry. *American Economic Review*, 112(6):1782–1817.

Franco, S. F. and Santos, C. D. (2021). The impact of Airbnb on residential property values and rents: Evidence from Portugal. *Regional Science and Urban Economics*, 88:103667.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5):1189–1232.

Garcia-López, M.-A., Jofre-Monseny, J., Martínez-Mazza, R., and Segú, M. (2020). Airbnb, short-term rentals and housing prices: Evidence from Barcelona. *Journal of Urban Economics*, 119:103278.

Gibbs, C., Guttentag, D., Gretzel, U., Morton, J., and Goodwill, A. (2018). Pricing in the sharing economy: a hedonic pricing model applied to Airbnb listings. *Journal of Travel & Tourism Marketing*, 35(1):46–56.

Glaeser, E. L., Gyourko, J., and Saks, R. E. (2006). Urban growth and housing supply. *Journal of Economic Geography*, 6(1):71–89.

Gotham, K. F. (2005). Tourism gentrification: The case of New Orleans' Vieux Carré (French Quarter). *Urban Studies*, 42(7):1099–1121.

Grisdale, S. (2021). Displacement by disruption: short-term rentals and the political economy of "belonging anywhere" in Toronto. *Urban Geography*, 42(5):654–680.

Gurran, N. and Phibbs, P. (2017). When tourists move in: How should urban planners respond to Airbnb? *Journal of the American Planning Association*, 83(1):80–92.

Guttentag, D. (2015). Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. *Current Issues in Tourism*, 18(12):1192–1217.

Gyourko, J., Mayer, C., and Sinai, T. (2013). Superstar cities. *American Economic Journal: Economic Policy*, 5(4):167–199.

Horn, K. and Merante, M. (2017). Is home sharing driving up rents? evidence from Airbnb in Boston. *Journal of Housing Economics*, 38:14–24.

Ke, Q. (2017). Short-term rentals and the sharing economy: The case of Airbnb. *International Journal of Hospitality Management*, 67:120–129.

Kelejian, H. H. and 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.

Kelejian, H. H. and Prucha, I. R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40(2):509–533.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.

Koenker, R. and Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives*, 15(4):143–156.

Koster, H. R. A., van Ommeren, J., and Volkhausen, N. (2021). Short-term rentals and the housing market: Quasi-experimental evidence from Airbnb in Los Angeles. *Journal of Urban Economics*, 124:103356.

Kuminoff, N. V., Parmeter, C. F., and 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.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74(2):132–157.

LeSage, J. P. and Pace, R. K. (2009). *Introduction to Spatial Econometrics*. CRC Press, Boca Raton, FL.

Li, H., Kim, Y., and Srinivasan, K. (2022). Market shifts in the sharing economy: The impact of Airbnb on housing rentals. *Management Science*, 68(11):8015–8044.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*, volume 30, pages 4765–4774.

Malpezzi, S. (2003). Hedonic pricing models: A selective and applied review. In O'Sullivan, T. and Gibb, K., editors, *Housing Economics and Public Policy*, pages 67–89. Blackwell Science, Oxford.

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.

Mullainathan, S. and Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.

Nieuwland, S. and van Melik, R. (2020). Regulating Airbnb: How cities deal with perceived negative externalities of short-term rentals. *Current Issues in Tourism*, 23(7):811–825.

Palmquist, R. B. (2005). Property value models. In Mäler, K.-G. and Vincent, J. R., editors, *Handbook of Environmental Economics*, volume 2, pages 763–819. Elsevier.

Parmeter, C. F., Henderson, D. J., and Kumbhakar, S. C. (2010). Applied nonparametric instrumental variables estimation. *Econometric Reviews*, 29(4):398–418.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55.

Saiz, A. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics*, 125(3):1253–1296.

Shabrina, Z., Arcaute, E., and Batty, M. (2022). Airbnb and its potential impact on the London housing market. *Urban Studies*, 59(1):197–221.

Sheppard, S. and Udell, A. (2016). Do Airbnb properties affect house prices? Working paper, Williams College Department of Economics.

Sirmans, G. S., Macpherson, D. A., and Zietz, E. N. (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature*, 13(1):1–44.

Sundararajan, A. (2016). *The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism*. MIT Press, Cambridge, MA.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

Valentin, M. (2021). Regulating short-term rental housing: Evidence from New Orleans. *Real Estate Economics*, 49(1):152–186.

Wachsmuth, D. and Weisler, A. (2018). Airbnb and the rent gap: Gentrification through the sharing economy. *Environment and Planning A: Economy and Space*, 50(6):1147–1170.

Wang, D. and Nicolau, J. L. (2017). Price determinants of sharing economy based accommodation rental: A study of listings from 33 cities on Airbnb.com. *International Journal of Hospitality Management*, 62:120–131.

Zervas, G., Proserpio, D., and Byers, J. W. (2017). The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. *Journal of Marketing Research*, 54(5):687–705.

Zietz, J., Zietz, E. N., and Sirmans, G. S. (2008). Determinants of house prices: A quantile regression approach. *Journal of Real Estate Finance and Economics*, 37(4):317–333.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.

A. Data Appendix

A.1. Variable Definitions

Table A1 provides a comprehensive list of variables used in the analysis, including their definitions, sources, and units.

A.2. City Name Standardisation

The Airbnb dataset contains city names with inconsistent formatting, including accented and unaccented variants and abbreviations. Table A2 lists the standardisation mapping applied during data cleaning.

A.3. Sample Attrition

The following summarises the sample attrition at each stage of data cleaning:

- **Airbnb:** Raw dataset ≈ 5,000 listings. After dropping listings with missing or non-positive prices, the cleaned sample retains 3,456 listings.
- **Rentals:** Raw dataset of 8,356 records. After restricting to single-family properties with monthly rental periods and positive parsed rents, the cleaned sample retains 8,303 listings, all with valid coordinates.
- **Spatial merge:** All rental listings with valid coordinates are retained; the Airbnb count is zero for listings with no Airbnb neighbours within the buffer radius (composition variables such as the mean nearby price are undefined in that case).
- **Regressions:** The baseline hedonic sample comprises the 7,925 rental listings with non-missing bedrooms and bathrooms.

B. Methodological Appendix

B.1. Haversine Distance Formula

The Haversine formula computes the great-circle distance between two points on the surface of a sphere, given their latitudes and longitudes. For two points (lat_1, lon_1) and (lat_2, lon_2) expressed in radians, the Haversine distance is:

$$d = 2R \cdot \arcsin\left(\sqrt{\sin^2\left(\frac{\Delta lat}{2}\right) + \cos(lat_1) \cdot \cos(lat_2) \cdot \sin^2\left(\frac{\Delta lon}{2}\right)}\right), \tag{A.1}$$

where:

- $R = 6{,}371$ km is the mean radius of the Earth;
- $\Delta lat = lat_2 - lat_1$ is the difference in latitudes;

Table A1. Variable Definitions

Variable	Source	Definition	Unit
Rental listing variables			
`rent`	Realtor.ca	Monthly asking rent	CAD
`log_rent`	Constructed	Natural log of monthly rent	—
`bedrooms`	Realtor.ca	Number of bedrooms	Count
`bathrooms`	Realtor.ca	Number of bathrooms (total)	Count
`building_type`	Realtor.ca	Building type (Apartment, House, Row/Townhouse)	Category
`size_interior`	Realtor.ca	Interior floor area	sq ft
`lat`, `lon`	Realtor.ca	Latitude and longitude of the listing	Degrees
`city`	Realtor.ca	City / municipality name	Category
Airbnb listing variables			
`price_numeric`	Airbnb	Nightly asking price	CAD
`log_price`	Constructed	Natural log of nightly price	—
`property_type`	Airbnb	Property type (Rental unit, House, Cabin/Chalet, Condo, Apartment)	Category
`rating`	Airbnb	Star rating (0–5 scale)	Numeric
`num_reviews`	Airbnb	Number of guest reviews	Count
`is_superhost`	Airbnb	Superhost status indicator	Binary
`is_guest_favorite`	Airbnb	Guest favourite designation	Binary
`pets_allowed`	Airbnb	Whether pets are allowed	Binary
`is_entire_home`	Constructed	Indicator for entire-home property types	Binary
`lat`, `lon`	Airbnb	Latitude and longitude of the listing	Degrees
`city`	Airbnb	City name (standardised)	Category
Airbnb exposure variables (constructed via spatial buffer merge)			
`airbnb_count_`r	Constructed	Number of Airbnb listings within radius r	Count
`airbnb_density_`r	Constructed	Airbnb count / buffer area (πr^2)	Per km^2
`mean_airbnb_price_`r	Constructed	Mean Airbnb nightly price within radius r	CAD
`share_entire_home_`r	Constructed	Share of entire-home Airbnb listings within r	Proportion
`mean_rating_`r	Constructed	Mean rating of Airbnb listings within r	Numeric
`superhost_share_`r	Constructed	Share of superhost listings within r	Proportion
City-level variables			
`airbnb_count_city`	Constructed	Total Airbnb listings in the city	Count
`mean_airbnb_price_city`	Constructed	City-level mean Airbnb nightly price	CAD
`share_entire_home_city`	Constructed	City-level share of entire-home listings	Proportion
`mean_rent_city`	Constructed	City-level mean monthly rent	CAD

Notes: Buffer radii $r \in \{250\text{m}, 500\text{m}, 1\text{km}, 2\text{km}\}$. All monetary values are in Canadian dollars (CAD). "Constructed" indicates variables derived from the raw data during the cleaning and merge stages.

Table A2. City Name Standardisation Mapping

Original Name(s)	**Standardised Name**
Montréal	Montreal
Québec City, Quebec City, Quebec	Québec
Levis	Lévis
Saint Come	Saint-Côme
Sainte Adele, Sainte-Adele, Ste-Adèle	Sainte-Adèle
Saint Sauveur, Saint-Sauveur-des-Monts	Saint-Sauveur
Ste-Agathe-des-Monts	Sainte-Agathe-des-Monts

- $\Delta lon = lon_2 - lon_1$ is the difference in longitudes.

The Haversine formula provides an accurate approximation for short to medium distances on the Earth's surface. For the distances relevant in our application (typically less than 5 km), the approximation error relative to the exact geodesic distance (Vincenty's formula) is negligible.

Implementation. In our spatial merge (Section 3.3), the Haversine distance is computed in vectorised form using NumPy. For each chunk of rental listings, we compute the full pairwise distance matrix between the chunk and all Airbnb listings, yielding an $(n_{\text{chunk}} \times n_{\text{Airbnb}})$ matrix of distances. Buffer membership is determined by comparing each element of this matrix against the threshold radius r.

B.2. Spatial Weights Construction

The spatial weights matrix $\mathbf{W}$ used in Model 4 (Equation 6) is a k-nearest-neighbour matrix with $k = 5$. For each rental listing i, the neighbourhood $\mathcal{N}_i$ is the set of the five rental listings closest to i in Euclidean coordinate space (excluding i itself):

$$w_{ik} = \begin{cases} \frac{1}{5} & \text{if } k \in \mathcal{N}_i, \\ 0 & \text{otherwise}, \end{cases} \tag{A.2}$$

so that the spatial lag $\sum_k w_{ik} \ln(\text{rent}_k)$ is the mean log rent of listing i's five nearest neighbours. Row-standardisation guarantees that every observation has a well-defined, equally weighted neighbourhood, avoiding the empty-neighbourhood problem that a fixed-radius definition would create in sparse rural areas.

Choice of k. The choice of $k = 5$ balances two considerations. A smaller k would capture only the most proximate listings, yielding a responsive but noisy spatial lag; a larger k would average over a broader area, reducing noise but blurring relevant spatial variation—and, in dense urban cores, five nearest neighbours typically lie within a few hundred metres.

Estimation. The inclusion of a spatial lag of the dependent variable as a regressor introduces a well-known simultaneity problem: $\sum_k w_{ik} \ln(\text{rent}_k)$ is correlated with η_i whenever the errors are spatially correlated, so OLS estimation of Equation 6 would be inconsistent. We therefore estimate the SAR model by the generalised-moments instrumental-variable estimator of Kelejian and Prucha (`GM_Lag` in `spreg`), which instruments the spatial lag with spatially lagged exogenous regressors, and the SEM by the corresponding `GM_Error` estimator (Anselin 1988; LeSage and Pace 2009).

B.3. Quantile Regression Estimation

The quantile regression model at quantile τ (Equation 7) is estimated by minimising the asymmetrically weighted sum of absolute residuals:

$$\hat{\boldsymbol{\beta}}_\tau = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \rho_\tau\big(\ln(\text{rent}_i) - \mathbf{X}_i'\boldsymbol{\beta}\big)\,, \tag{A.3}$$

where $\rho_\tau(u) = u(\tau - \mathbf{1}[u < 0])$ is the check function (Koenker and Bassett 1978). We estimate the model with the iteratively reweighted least squares algorithm implemented in `statsmodels`.

Standard errors are the asymptotic estimates based on the kernel estimate of the conditional density of the response at the fitted quantile (the default in `statsmodels`), which are consistent under independent but not necessarily identically distributed errors.

B.4. Machine-Learning Model Details

LASSO and Elastic Net. The LASSO model (Tibshirani 1996) estimates:

$$\hat{\boldsymbol{\beta}}_{\text{LASSO}} = \arg\min_{\boldsymbol{\beta}} \left\{ \frac{1}{2n} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_1 \right\}, \tag{A.4}$$

where $\lambda > 0$ is the regularisation parameter selected by 5-fold cross-validation. The elastic net extends this with an L_2 penalty, controlled by a mixing parameter $\alpha \in (0, 1)$.

Random Forest. The random forest model (Breiman 2001) constructs an ensemble of $B = 500$ regression trees with a maximum depth of 15, each trained on a bootstrap sample with random feature subsampling.

Gradient Boosting. The gradient boosting model (Friedman 2001) sequentially fits shallow regression trees to the residuals of the current ensemble. We use 500 boosting iterations, a maximum tree depth of 5, and a learning rate of 0.05. SHAP (SHapley Additive exPlanations) values (Lundberg and Lee 2017) are computed on the held-out test set from the fitted gradient boosting model to quantify the contribution of each feature to individual predictions.

C. Additional Robustness Results

This appendix presents additional robustness checks that supplement the results reported in Section 6.

C.1. Full Regression Output — Baseline Hedonic Model

The preferred hedonic rent specification (Model 1c in Table 3) includes property controls (bedrooms, bathrooms, building type) and city fixed effects. Full regression output including all control variable coefficients and city fixed effects is available upon request.

C.2. Correlation Matrix of Airbnb Exposure Measures

The pairwise correlation matrix of the Airbnb exposure variables at the 500 m buffer radius documents the degree of collinearity among alternative exposure measures. The various exposure measures—count, density, share of entire-home listings, mean price—are positively correlated, reflecting the spatial concentration of Airbnb activity.

C.3. Quantile Regression — Full Coefficient Tables

The quantile regression results reported in Table 7 focus on the Airbnb exposure coefficient across the rent distribution. Full coefficient tables including control variables at each quantile are available upon request.

C.4. Machine-Learning Cross-Validation Results

Table 8 in the main text reports hold-out sample performance for all machine-learning models. The gradient boosting and random forest models achieve substantially higher R^2 on the test set relative to linear models, suggesting meaningful nonlinearities in the rent–amenity relationship.

C.5. Additional Subsample Results

The subsample analysis in Section 6 examines Montreal versus non-Montreal and apartment versus house subsamples. Additional subsample results (e.g., by Airbnb exposure quartile or by borough-level regulation status) are available upon request.