Parcel-Level Land Valuation under Planning Policy Change:

A Spatially Based and Segmented Modeling Framework

by

Zekai Wu

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Master of Applied Science

in

Civil Engineering

Waterloo, Ontario, Canada, 2026

© Zekai Wu 2026

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis was supported through a collaborative research effort with LandLogic Inc. LandLogic™ is a Land Modernization Platform powered by AI which provides automated land due diligence reports quickly.

Dr. Carl Haas and Dr. Amin Sarang provided high-level guidance on the overall research direction and contributed expertise in machine learning concepts and analytical insight throughout the development of the study. Dr. Haas guided crafting of the thesis document through several iterations. Dr. Arash Shahi offered high-level directional guidance and access to LandLogic resources used in this research. Dr. Amin Sarang also contributed insights on planning practices and offered guidance on how to interpret and apply urban planning frameworks relevant to land valuation analysis. Saeid Emamgholian provided guidance on GIS-related questions and, together with Soroosh Sanjarani and LandLogic teams, coordinated the provision and organization of the LandLogic datasets used in this research.

Abstract

Planning policies shape land markets by defining development rights, regulating land use intensity, and signaling future growth expectations. However, in current practice, land and parcel value estimation remains highly dependent on expert judgment and manual comparison of recent transactions. While professional appraisal methods are effective for individual assessments, they rely heavily on subjective interpretation and lack systematic mechanisms to identify how planning policy changes influence parcel values across large datasets. This limitation is particularly evident in growing suburban municipalities, where frequent policy updates are used to respond to development pressure and evolving growth objectives. As a result, both municipal decision-making and real estate analysis increasingly require data-driven models capable of automatically evaluating the valuation impacts of planning policy changes.

This thesis addresses this gap by developing a spatially explicit, data-driven framework to examine how planning policy changes are capitalized into parcel-level land values in the Town of Aurora, Ontario. The research integrates GIS-based spatial analysis with segmented regression and machine learning modeling to move beyond manual valuation approaches. Parcel-level datasets are constructed by combining transaction records with Official Plan designations, zoning regulations, and adjacent based spatial variables. Parcels are further segmented by size into two datasets to distinguish between house driven and land driven valuation mechanisms, enabling the model to identify the conditions under which planning signals emerge in observed prices.

The findings indicate that parcel size fundamentally mediates valuation behavior. For small residential parcels, price patterns are dominated by existing structures, obscuring planning policy effects. In contrast, for larger parcels with redevelopment potential, regulatory designations, environmental constraints, and spatial context are systematically reflected in land value. By replacing purely expert-driven inference with an interpretable and replicable modeling approach, this thesis enhances the transparency of parcel value estimation and strengthens the empirical evaluation of planning policy impacts in suburban land markets.

Acknowledgements

I would like to gratefully acknowledge the financial and research support provided by Mitacs and LandLogic™ . Their commitment to industry-academic collaboration provided me with a valuable opportunity to conduct applied research at the intersection of planning, GIS, AI, and land valuation.

I sincerely thank my supervisors, Dr. Carl Haas and Dr. Amin Sarang, for their guidance, encouragement, and support throughout this research. Their mentorship has been instrumental in both my academic and personal development.

I would also like to thank Dr. Arash Shahi, Dr. Saeid Emamgholian, Soroosh Sanjarani, and the entire LandLogic™ team for their support and providing access to the data and resources that made this research possible.

My sincere appreciation is extended to Dr. Joyce Kim and Dr. Nadine Ibrahim for serving on my thesis committee and for thoughtful review and value feedback on this work.

Finally, I am deeply grateful to my family for their warm encouragement, patience, and support.

I hope that this research contributes, in some small way, to more informed planning decisions and to the broader goal of sustainable development for the benefit of our shared Earth.

Table of Contents

Author's Declaration ii

Statement of Contributions iii

Abstract iv

Acknowledgements vi

List of Figures xi

List of Tables xii

List of Abbreviations xiii

Chapter 1 Introduction 1

1.1 Problem Statement 2

1.2 Background 5

1.3 Research Objectives and Contribution 7

1.4 Scope 8

1.5 Methodology 9

1.6 Thesis Organization 11

Chapter 2 Literature Review 13

2.1 Conceptual Foundations: Planning Policy Instruments and Land Valuations 16

2.1.1 The Strategic Role of Land Use in Shaping Land Value Expectations 16

2.1.2 The Economic Function of Zoning Regulations in the Land Market 18

2.1.3 Integrating Planning Regulations into Land Valuation Analysis 19

2.2 Conceptual and Technical Foundations of Land Valuation Models 20

2.2.1 Mechanisms for Integrating Real Estate and Spatial Data in Land Valuation 21

2.2.2 Economic and Financial Modeling Approaches for Estimating Land Value 22

2.3 Key Theoretical Questions and Scholarly Debates 23

2.3.1 Core Questions on Planning Interventions and Land Markets 24

2.3.2 Conflicting Findings and Methodological Challenges 25

2.4 Identified Research Gap: Parcel Level Regulatory Analysis in Outlying Communities such as Aurora within Large Urban Areas such as the GTA (Greater Toronto Area) 26

2.4.1 Limitations of Current Literature 27

2.4.2 Justification for Localized, Spatial-Policy-Integrated Valuation 29

2.5 Concept Models: Regression and Prediction Modeling 29

2.5.1 Multi-Regression 30

2.5.2 Prediction Modeling 32

Chapter 3 Proposed Methodology 34

3.1 Data Collection and Preparation 36

3.2 Variable Selection 39

3.2.1 Data Preparation and Variable Selection 39

3.2.1.1 Physical Factors 44

3.2.1.2 Legal and Planning Controls 45

3.2.1.3 Economic Factors 47

3.2.1.4 Adjacency Characteristics 48

3.2.2 Model Specification 49

3.3 Regression Modeling Approach 50

3.3.1 Model Structure and Rationale 50

3.3.2 Segmented Regression by Parcel Size 51

3.3.3 Variable Transformation and Scaling 54

3.3.4 Model Identification and Validation in Machine Leaning Model Light GBM 55

Chapter 4 Results 57

4.1 Segmented Regression Results 58

4.1.1 Model A: Parcels < 0.3 Acre (House-Driven Value) 58

4.1.2 Model B: Parcels ≥ 0.3 Acre (Land-Driven Market) 64

4.2 Machine Leaning Model Light GBM 71

4.3 Comparative Discussion of the Two Segmented Regression Models and the Machine Leaning Model made with LightGBM 74

4.3.1 Comparison of Model Performance 75

4.3.2 Comparison of Variable Influence 75

4.3.3 Integrated Interpretation 77

Chapter 5 Interpretation and Discussion 79

5.1 Valuation of Planning Policy Change's Impact on Land Price 79

5.1.1 Contribution to Digital Planning Platform and City Service in Aurora 80

5.1.2 Generalization to Other Municipalities and Secondary Benefits 81

5.2 Limitations 82

5.3 Further Implications 84

Chapter 6 Conclusions and Recommendations 86

6.1 Summary of Findings 86

6.2 Summary of Research Contribution 87

6.3 Future Research 88

6.4 Conclusions 90

References 93

List of Figures

Figure 1 OP 2010 vs OP 2024 affected parcels 4

Figure 2 Distribution of Aurora's Parcel Sizes 4

Figure 3 Town of Aurora's Location 7

Figure 4 Research Framework and Analytical Workflow 11

Figure 5 Overall Research Variables and Methodological Framework 35

Figure 6 Parcel Size Distribution 53

Figure 8 Normal P-P Plot of Regression Standardized Residual from SPSS of Model A: Parcels < 0.3 Acre (House-Driven Value) 62

Figure 10 Normal P-P Plot of Regression Standardized Residual from SPSS of Model B: Parcels ≥ 0.3 Acre (Land-Driven Market) 69

Figure 11 Distribution of Prediction Residuals from the LightGBM Model 73

Figure 12 Comparison of Predicted and Observed Parcel Values Using the LightGBM Model ... 73

List of Tables

Table 1 Literatures and This Thesis Comparison on Planning Policy Focus, Spatial Scale, Methodology, and Contribution 14

Table 2 LandLogic™ Source Data Attribute..... 38

Table 3 Dependent Variables 40

Table 4 Independent Variables..... 41

Table 5 Diagnostic Evaluation of Model Residual 59

Table 6 Coefficients from SPSS of Model A: Parcels < 0.3 Acre (House-Driven Value) 61

Table 7 Coefficients from SPSS of Model B: Parcels ≥ 0.3 Acre (Land-Driven Market) 68

Table 8 LightGBM Model Performance Metrics for Parcel Value Prediction 73

List of Abbreviations

Abbreviation	Definition
AVMs	Automated Valuation Models
BWM	Best-Worst Method
OP	Official Plan
ZBL	Zoning By-Law
GTA	Great Toronto Area
The Act	The Ontario Planning Act
MRA	Multiple Regression Analysis
HPI	Home Price Index
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
RMSLE	Root Mean Squared Log Error
MAPE	Mean Absolute Percentage Error

Chapter 1 Introduction

Land valuation plays an essential role in planning, real estate markets, and municipal investing. In rapidly growing regions, understanding how planning policies influence land price is critical for ensuring transparency, fairness, and efficient land markets. While traditional valuation methods have relied heavily on comparable market transaction and professional manual judgment, these methods predominantly estimate property price, where the observed price reflects the combined influence of land and existing structures. Emerging research increasingly emphasizes the integration of spatial data and empirical modeling to capture the dynamic effects of planning interventions on parcel value, which includes both land driven and house driven value (Kauko & d'Amato, 2009; MPAC, 2025; Wyatt, 2022). This shift reflects a growing recognition that parcel prices are shaped not only by inherent physical and locational attributes but also by: regulatory frameworks; land use designation; parcel size which drives relative land and house impacts on value; and policy changes over time. Recent advances in parcel-level datasets, spatial analytics, and machine-learning models allow researchers and planners to examine these relationships with greater precision, particularly in the context where planning policy changes may produce differentiated market responses across parcel types. Therefore, as communities face increasing development pressures and evolving policy goals, there is a pressing need for robust valuation approaches that can systematically account for these factors and inform evidence-based decision-making (Demetriou, 2016). This need is particularly rising in municipalities such as the Town of Aurora, where an Official Plan had undergone a review and

changes made on land use designation, creating a complex valuation environment that requires a fast automatic land valuation by or for those with land interests.

1.1 Problem Statement

In Ontario's fast-changing land use regime, particularly in municipalities such as the Town of Aurora (Aurora)(Joseph, 2025), the municipal planners and private developers face a critical gap in their ability to systematically and automatically estimate land value. Despite the increasing availability of digital tools and planning data, land valuation in these jurisdictions remains largely dependent on professional intuition and manual case-by-case assessment. There is currently no widely adopted method that allows planners or policymakers to quantify, in advance, how proposed changes to Official Plan (OP) or Zoning By-Law (ZBL) designations will impact land prices at the parcel level. In other words, this limitation becomes increasingly significant as municipalities adopt more complex planning policy frameworks, where the price effects of regulatory change may vary substantially across parcel types, development capacities, and spatial contexts.

This absence of impact information leads to significant limitations for evidence-based planning and development decision-making. Without a strong, spatially sensitive method to estimate how planning interventions influence land value, municipalities are constrained in their ability to anticipate the market implications of land use policy, guide infrastructure investment, or align development engagements with planning goals. Similarly, developers and real estate agents often lack a consistent framework to assess the profitability of land acquisition under changing

regulatory scenarios (Robinson et al., 2021). Planning of this capacity is particularly problematic in contexts like Aurora, where land use policy evolved significantly from 2010 to 2024, and where planning decisions must balance growth management, intensification, and environmental protection within a highly dynamic real estate market. The challenge is acute in Aurora, where more than 70 percent of parcels were impacted by the OP updates, and more than 90 percent of parcels are under 0.3 Acres (See Figure 1 and Figure 2), which are subject to single home (house) structure driven valuation. This creates a valuation environment in which planning signals can be obscured by structure-driven pricing, making it difficult to interpret policy effects on housing and land values, separately or combined.

This thesis addresses this gap by developing an approach that integrates OP, ZBL, and real estate data into parcel-level land valuation analysis, thereby offering a more transparent, replicable, and policy-sensitive method for understanding how planning decisions shape land value outcomes. Accurate automatic estimation of land valuation is thus provided, potentially speeding up the housing development process. The proposed approach offers a transparent, replicable, and policy-responsive method for estimating land value across heterogeneous parcel types, thereby improving the capacity of municipalities and private-sector actors to understand, anticipate, and respond to planning-policy impacts. By advancing a more systematic and automated valuation method, the study contributes to the broader goal of accelerating development review processes and supporting data-driven decision-making in Ontario's rapidly evolving land markets.

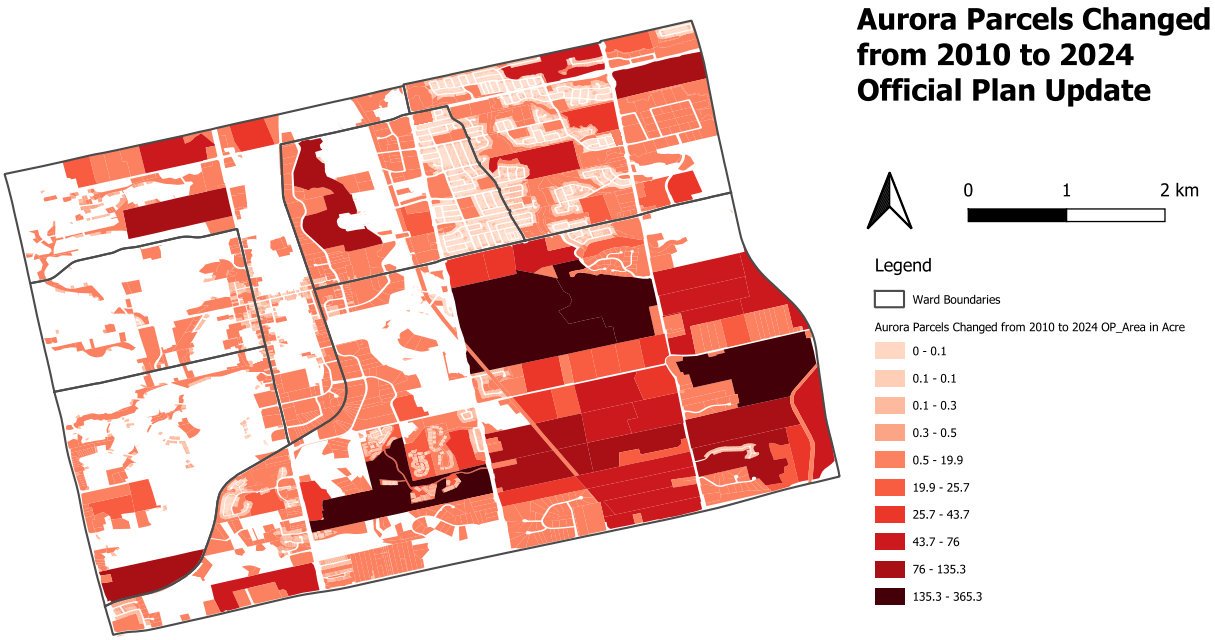


Figure 1 OP 2010 vs OP 2024 affected parcels

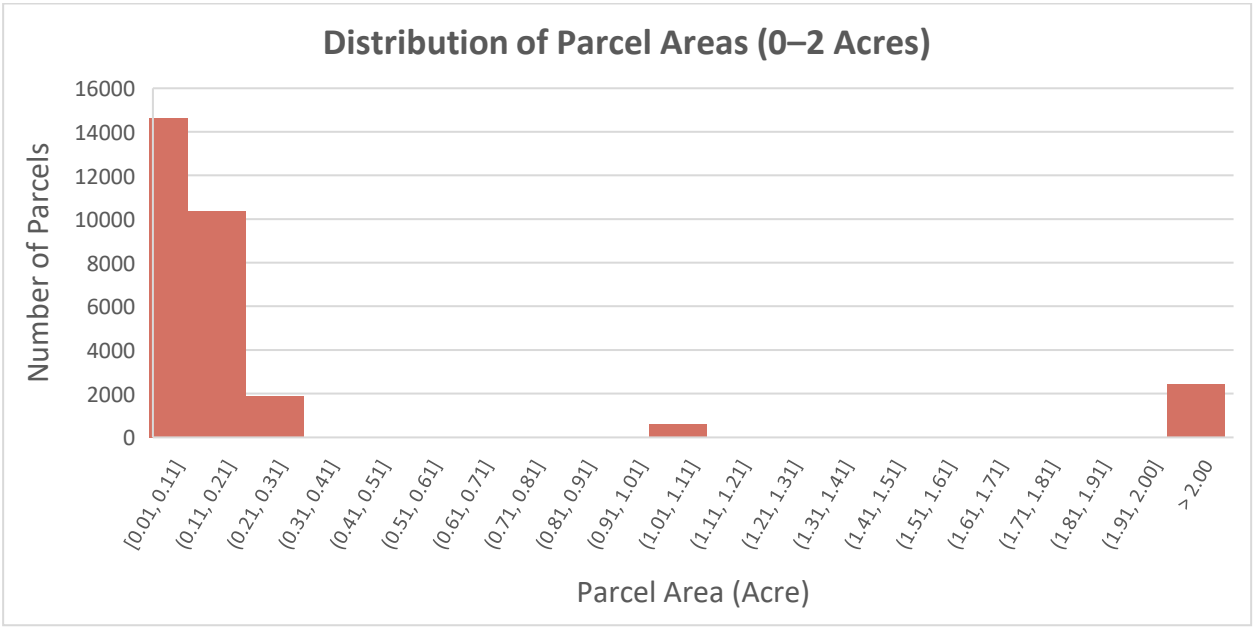


Figure 2 Distribution of Aurora's Parcel Sizes

1.2 Background

Land valuation has long been a major concern in planning, real estate economics, and municipal finance. Accurate estimation of land value supports a wide range of planning activities, including development review, infrastructure financing, land-use regulation, and growth management (Stewart, 2022). However, the traditional valuation approaches have relied primarily on market transactions, professional appraisal practices, and localized expertise (Ontario Professional Planner Institute, 2018). Where these methods remain essential in practice, in European cities similar to Aurora, they often exhibit lack of transparency, efficiency, and scalability, particularly in rapidly evolving planning contexts where regulatory frameworks and land-use policies are subject to frequent change (Hersperger et al., 2022). As municipalities increasingly adopt data-driven and digital planning tools, there is growing interest in analytical approaches that can systematically evaluate how planning policies interact with spatial characteristics to shape land-value outcomes.

The Town of Aurora provides an appropriate and policy-relevant case study for examining these dynamics. Aurora is located in the central-northern portion of the Greater Toronto Area (See Figure 3), within York Region, and functions as a rapidly growing suburban municipality influenced by regional growth pressures from the Toronto metropolitan area. Its strategic location along major transportation corridors, including Highway 404 and the GO Transit network (also known as the Major Transit Station Area, MTSA), has positioned Aurora as an attractive location for residential development, employment growth, and intensification initiatives (Town of Aurora, 2026b). As a result, the municipality has experienced sustained

population growth and increasing development activity over the past two decades (Town of Aurora, 2026a).

Aurora's land-use structure reflects a predominantly low-density residential urban form, characterized by a large proportion of small residential parcels, alongside a more limited number of larger parcels designated for employment, commercial, or mixed-use development. While the majority of parcels are single-to-multifamily-structure-driven residential lots, much of the municipality's future growth and intensification potential is concentrated in a smaller subset of development-scale parcels. These parcels are often subject to more complex planning-policy frameworks, including OP land-use designations, zoning density permissions, Major Transit Station Area policies, and environmental protection overlays (Town of Aurora, 2026b).

In 2024, Aurora had undertaken significant planning-policy updates, including revisions to its OP to align with provincial growth policies and regional planning objectives. These changes aim to balance intensification, environmental protection, and housing supply within a constrained land base (Town of Aurora, 2025). However, the extent to which such planning-policy changes are capitalized into land prices at the parcel level remains poorly understood, particularly given the municipality's heterogeneous parcel-size distribution and mixed land-use structure.

Against this backdrop, Aurora offers a compelling context for investigating how planning-policy instruments interact with spatial characteristics and parcel attributes to influence land value. Its combination of rapid growth, evolving planning frameworks, and contrasting parcel-scale conditions make it an ideal setting for developing and testing a parcel-level, policy-sensitive land

valuation framework that can inform both municipal planning practice and broader academic inquiry.

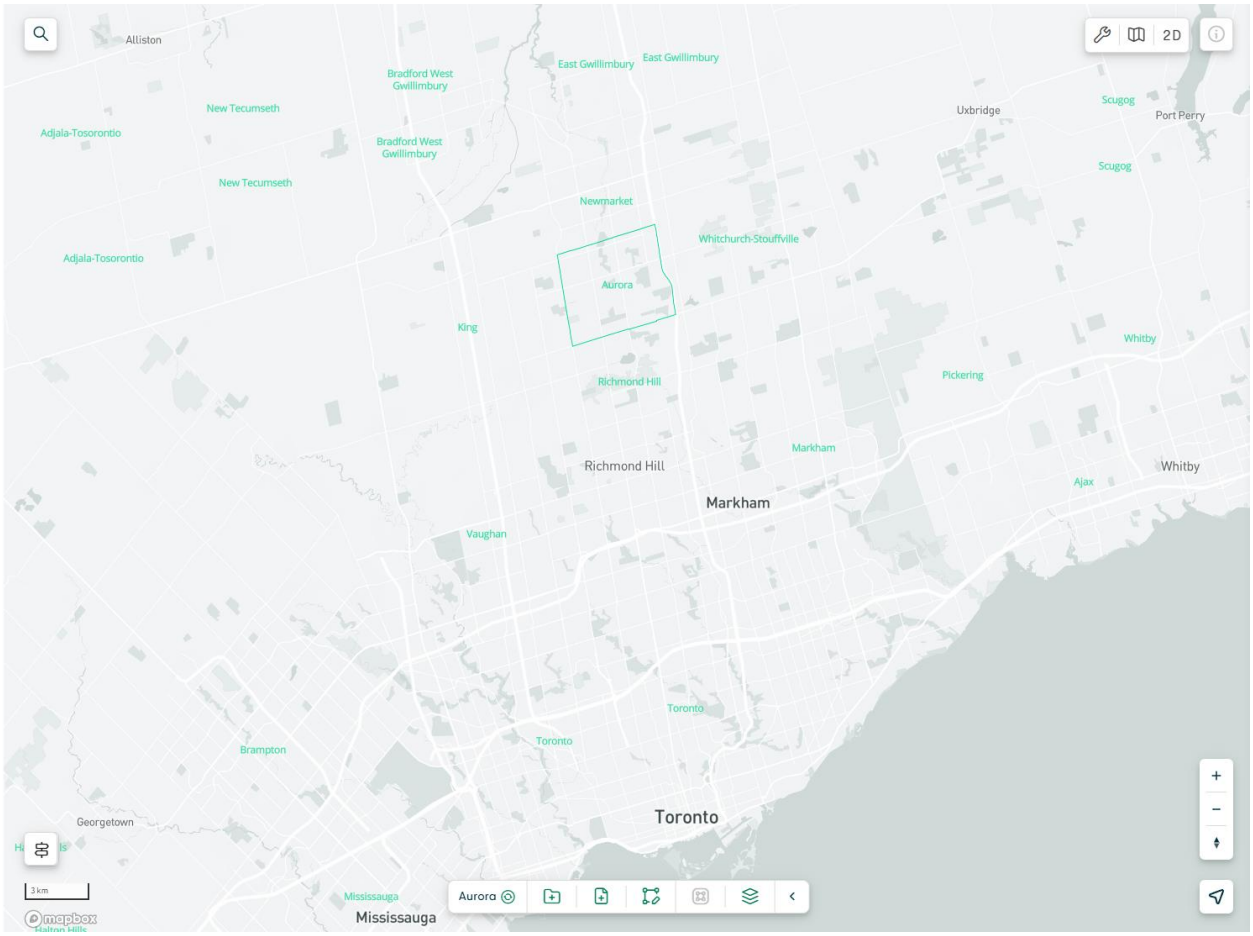


Figure 3 Town of Aurora's Location

1.3 Research Objectives and Contribution

The primary objective of this research is to examine how planning policies including land use and zoning changes influence land valuation at both the parcel and municipal scale. Specifically, this study aims to find a model to facilitate the calculation of land value with the following steps:

- Analyze how different physical and economic factors correlate with land value;
- Evaluate how adjacent land use factors influence the value of a target parcel;
- Assess how land use designation changes between 2015 and 2024 Official Plans affect land value;
- Apply these insights in a simplified model to identify patterns in land value change.

By doing so, the research contributes to planning and valuation practice by developing a spatial-policy framework that explains how and why land value is produced across space, not just within individual parcels.

1.4 Scope

This study focuses on the Town of Aurora, Ontario, and draws on transaction data from 2014 to 2025, zoning and OP data, and development application records provided by LandLogic Inc. The scope includes:

- All parcels with valid land transaction records within the period between 2014 to 2025.
- Official zoning and land use designations as per 2015 and proposed 2024 Official Plans.
- Adjacent land use composition defined as the land use designation of parcels that directly share a boundary with the target parcel. (Adjacent parcels are identified through direct polygon-to-polygon contact, excluding parcels that are separated by roads, rights-of-way, or buffer distances.)

Within this study, “land” refers to the parcel as a spatial and regulatory unit, characterized by its size, planning designations, zoning permissions, and immediate regulatory context, rather than

by existing building structures. The analysis does not cover building-level characteristics or income-based valuation metrics and is limited to land value per acre.

1.5 Methodology

This research applies an integrated spatial and empirical methodology to examine how planning-policy designations, adjacency conditions, and regulatory constraints influence parcel-level land values in Aurora. The methodological framework is designed to support transparent and replicable land valuation analysis using planning and GIS data.

The analysis begins with a comparison of land values across zoning and Official Plan designation categories to establish baseline relationships between planning designations and land price per acre. It then examines adjacency effects by evaluating how the composition of nearby land uses influences parcel value. These spatial relationships are quantified using GIS-based proximity measures to capture neighborhood-level land-use interactions.

The methodology further evaluates planning-policy change by comparing parcels with different Official Plan designations between 2010 and 2024, allowing the study to assess how policy transitions relate to observed valuation patterns. Physical and infrastructural characteristics, including ward location and access to the transportation network, are incorporated to account for spatial heterogeneity. Regulatory constraints are evaluated through measures of environmentally restricted land that require additional approvals and may limit development potential.

To separate the difference between house driven and land driven valuation dynamic, the analysis proposes a segmented regression framework that separates parcels by size. Smaller parcels are more likely to be priced on physical housing characteristics and structure, reflecting house driven valuation behavior (akin to housing property values). In contrast, larger parcels are more likely to exhibit land driven valuation patterns, where planning policies, development opportunities, and regulatory constraints have greater influence on observed price. This segmentation allows the model to isolate the conditions under which planning signals are capitalized into parcel value.

Spatial analysis is conducted using GIS, with supplementary land-use identification performed through LandLogic.ai. Data processing and regression analysis are implemented using Excel and IBM SPSS. Together, these methods provide a policy-sensitive and data-driven framework for assessing the impact of planning decisions on land value.

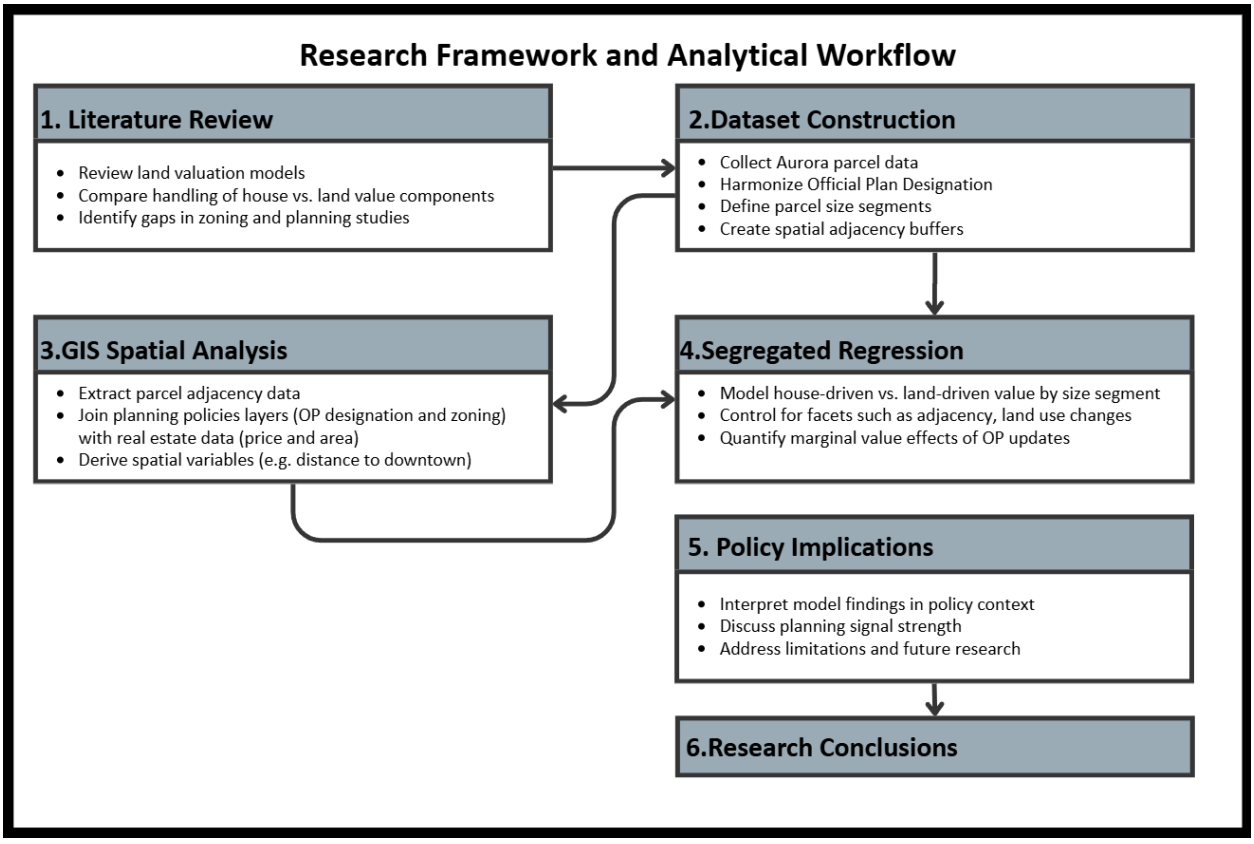


Figure 4 Research Framework and Analytical Workflow

1.6 Thesis Organization

This thesis is structured as follows:

- Chapter 1 introduces the problem, objectives, and scope of the research.
- Chapter 2 reviews literature on land valuation, adjacency effects, zoning theory, and Official Plan implementation. Discussed methods and modules used in land valuation.
- Chapter 3 presents the data sources and introduces methodologies used in land valuation.
- Chapter 4 discussed the outcomes of the research.

- Chapter 5 covers interpretation and discussion of research's value, limitations, and further implications.
- Chapter 6 concludes with findings, contribution of the research, and suggestions for future research.

Chapter 2 Literature Review

Building on the recognition that land prices reflect a complex interplay between market dynamics and regulatory interventions, this chapter reviews existing research on conceptual foundations, empirical evidence, and methodological approaches to land valuation (See Table 1). While traditional appraisal practices typically focus on comparable sales and professional judgment (Glumac & Des Rosiers, 2021), more recent studies highlight the importance of integrating spatial data, zoning designations, and official land use policies to better capture how planning decisions shape property values over time (Ayalke & Sisman, 2022; Gunes & Yomralioglu, 2019). Hedonic pricing models have emerged as a widely used framework for decomposing land value into constituent attributes (Kok et al., 2014), while automated valuation models (AVMs) and machine learning techniques are increasingly applied to enable scalable data-driven assessments (Khalid, 2023). Despite these advances, limited attention has been given to combining parcel-level zoning, land use, and adjacent land characteristics into integrated valuation models, particularly in the context of suburban municipalities undergoing rapid policy change. Accordingly, this literature review synthesizes theoretical perspectives and empirical findings to establish a structured literature map, clarify how planning policies are operationalized in valuation models, and derive the variable selection and modeling logic that guide the spatially explicit valuation approach adopted in this thesis.

Table 1 Literatures and This Thesis Comparison on Planning Policy Focus, Spatial Scale, Methodology, and Contribution

Study	Policy Focus	Spatial Scale	Core variables emphasized	Methodology approach	Key contribution	Limitation addressed by this thesis
Mandelker (1976)	Official Plans and zoning hierarchy	Legal system	Plan conformity, regulatory hierarchy	Legal analysis	Establishes OP as superior planning instrument guiding zoning	Does not quantify land value effects
Jaeger (2006)	Land use regulation and exemptions	Parcel	Regulatory constraint vs permit value	Conceptual economic analysis	Distinguishes policy effect from site-specific approval value	Lacks empirical valuation modeling
Han et al. (2020)	Land use regulation	City and parcel	Density control, land supply restriction	Panel regression	Shows land use control capitalized into prices	Limited spatial adjacency treatment
Freemark (2019)	Zoning reform	Parcel	FAR, parking minimums, density	Difference-in-differences	Demonstrates rapid price capitalization after upzoning	Focuses on large urban context
Kendall & Tulip (2018)	Zoning restrictions	City	Zoning premium, land scarcity	Hedonic regression	Separates land and structure driven pricing	No parcel-level adjacency modeling
Irwin & Bockstael (2003)	Land use and development spillovers	Parcel	Adjacent land use, development status	Spatial hazard modeling	Identifies spatial externalities in land markets	Does not model land value explicitly
Ayalke & Sisman (2022)	Planning and spatial constraints	Parcel	Distance, land use, environmental layers	GIS + MCDA	Demonstrates valuation without transactions	Subjective weighting of factors
Gunes & Yomralioglu (2019)	Land use and infrastructure	Parcel	Zoning, accessibility, services	GIS-based mathematical appraisal	Enables transparent spatial valuation	Limited market validation
Jafary et al. (2024)	Zoning and land use in AVMs	Multi-scale	Zoning codes, density allowances	AVM review	Establishes regulatory data as key AVM input	Limited policy interpretation
Kok et al. (2014)	Land use regulation	City	Regulation intensity, land price	GMM panel regression	Identifies scarcity and signaling effects	Aggregated spatial resolution
Ecker & Isakson (2005)	Parcel size effect	Parcel	Lot size nonlinearity	Hedonic regression	Identifies structure-dominated size threshold	Does not integrate planning variables

Clauretie & Li (2017)	Development-scale parcels	Parcel	Price per acre, scale effects	Regression and predictive models	Shows parcel size dependent valuation regimes	Limited policy integration
This thesis	OPs, zoning, adjacency lands	Parcel	OP change, zoning permission, adjacent land shares	GIS + segmented regression	Integrates planning policy and spatial context in valuation	Addresses suburban small-parcel bias

2.1 Conceptual Foundations: Planning Policy Instruments and Land Valuations

The spatial structure and economic performance of urban land markets are heavily influenced by the planning policy tools that govern development permissions and growth management. In particular, OPs and ZBLs are two significant regulatory tools that heavily influence land supply, use intensity, and parcel value.

This section examines the conceptual foundation for planning-influenced land valuation. Subsection 2.1.1 explores how land use designations in OPs influence land value expectations by controlling land uses and signaling growth potential. Subsection 2.1.2 investigates the economic implications of zoning, particularly how regulatory changes affect market pricing using empirical cases such as the upzoning reforms in Chicago. Finally, Subsection 2.1.3 discusses the integration of both OPs and zoning by-laws into contemporary land valuation methods. Recent advances are emphasized in automated valuation modeling and the growing recognition of spatial regulatory context as an essential input.

2.1.1 The Strategic Role of Land Use in Shaping Land Value Expectations

Municipal Official Plans (OPs) are also known as community plans or comprehensive plans in other provinces or countries. They define long-term land use policy and suggestions by designating areas for specific uses, which follow the development goals and proposes of the local government. OPs are designated to be periodically reviewed and updated in response to evolving growth needs. Under the Ontario Planning Act (the Act), municipal councils are required to review their OPs at least once every five years to ensure consistency with provincial

policies and the changing development conditions (Hodge et al., 2014). Thus, the changes' impact on land use plays a significant role in influencing land market by shaping what kind of uses are permitted and anticipated.

Research shows that changes in land use policy, especially through OP redesignation, are often capitalized into land price. Typically, the higher permitted densities under land use regulations are positively and significantly capitalized into land prices. Notably, the authors highlight that residential and commercial land supply controls also raise land values, while administrative quotas on construction land appear ineffective(Han et al., 2020).

Further conceptual clarity is offered by Jaeger (2006), who distinguishes between the general policy effect of land use regulation and the site-specific value of exemptions. Using the case of Oregon's Measure 37, he argues that the value added by individual development permits or waivers does not prove that the regulation itself reduced value. This distinction is highly relevant in the context of Aurora, where approvals to develop within restrictive areas, such as lands under the Oak Ridges Moraine Conservation Plan, may elevate parcel value. However, these gains often stem from exclusive development access rather than a net reduction in market value caused by the regulation(Jaeger, 2006).

Mandelker (1976) emphasizes that zoning policies must derive from and conform to comprehensive plans (which are similar to OPs). His analysis of U.S. case law confirms the legal and functional hierarchy of planning instruments, where the OP sets the structure for subordinate zoning controls (Mandelker, 1976). This principle is also applicable in the Ontario planning framework. The land use types are primarily controlled by OPs, while ZBLs provide

more detailed controls such as density regulation, development requirements, and permitted uses.

2.1.2 The Economic Function of Zoning Regulations in the Land Market

Zoning by-laws (ZBLs) are detailed legal instruments that assign land into zones with specified permitted uses, densities, and development requirements. In contrast to the OP's strategic scope, zoning functions at the parcel level and directly determines land use rights. Because zoning controls development capacity and form, it has an immediate and measurable influence on land prices. This section explores how zoning, especially in cases of regulatory change, affects land values.

The empirical study by Freemark (2019) on Chicago's 2013 and 2015 upzoning reforms offers a comprehensive analysis of zoning's impact on property valuation. By applying a difference-in-differences method to property transaction data, Freemark finds that parcels receiving increased density allowances, such as higher floor-area ratios and reduced parking minimums, experienced significant price increases of 15 to 23 percent within two years of the reform. The study found that zoning changes led to price shifts, which suggests that zoning influences land value expectations quickly (Freemark, 2019).

This observation is further supported by Kendall and Tulip (2018), who use hedonic regression models to demonstrate that zoning restrictions account for a substantial portion of housing price inflation across Australian cities. By separating land values from construction costs, they estimate zoning premiums that reflect constrained development capacity and increased

competition for limited zoned parcels. Their findings reinforce the importance of treating zoning as a core variable in any analysis of land valuation(Kendall & Tulip, 2018).

Collectively, these studies highlight that zoning regulations serve as more than development controls; they function as price-determining mechanisms in regulated land markets. As such, zoning must be treated not as a static background variable but as a dynamic and localized determinant of land value.

2.1.3 Integrating Planning Regulations into Land Valuation Analysis

Recent developments in land valuation methodology emphasize the integration of planning data, especially zoning and land use policies into predictive models. Automated Valuation Models (AVMs), which use machine learning and spatial analysis to estimate land value, increasingly rely on regulatory information to improve accuracy.

Jafary, Shojaei, Rajabifard, and Ngo (2024) conduct a systematic review of AVM development, finding that zoning codes, land use types, and density allowances are essential inputs for reliable land price estimation. Their review underscores the shift toward planning-aware valuation, where the inclusion of regulatory data enhances model precision and contextual fit. In regulated jurisdictions like Aurora, AVMs that incorporate zoning and OP inputs are better positioned to reflect actual market behavior(Jafary et al., 2024).

The earlier work by Irwin and Bockstael (2002) complement this perspective by incorporating spatial spillovers into land use change modeling. Their parcel-level analysis in Maryland reveals that development decisions are influenced not only by a parcel's zoning but also by the

development status of neighboring parcels. Their spatial hazard model shows that development is less likely in areas already surrounded by residential land, indicating the presence of negative spatial externalities. These insights are particularly relevant to this thesis, which aims to quantify the influence of both zoning and adjacent land use context on land valuation in Aurora (Irwin & Bockstael, 2003).

Together, these studies establish that land valuation models must account for the regulatory environment as well as the spatial and market dynamics in which parcels are embedded. For Aurora, this means building a valuation framework that treats zoning permissions, OP designations, and adjacent parcel attributes as integral components—not as peripheral data points. Integrating these elements into AVM development ensures a more accurate and policy-sensitive approach to land valuation.

2.2 Conceptual and Technical Foundations of Land Valuation Models

Land valuation is a multifaceted process that requires both spatial and economic perspectives to capture the value embedded in location, policy context, and development potential. This section synthesizes two essential streams of literature: one focuses on the technical integration of real estate and geospatial data into valuation workflows, and the other addresses the economic and financial modeling techniques used to estimate land values. Together, they form a comprehensive conceptual foundation for data-driven land valuation, particularly under changing planning policy conditions.

2.2.1 Mechanisms for Integrating Real Estate and Spatial Data in Land Valuation

A growing body of literature emphasizes the importance of integrating spatial datasets, such as zoning boundaries, land use designations, proximity to amenities, and environmental constraints—into land valuation models. GIS-based frameworks enable researchers and planners to analyze the spatial distribution of land value drivers and conduct mass appraisals.

For example, Ayalke and Sisman (2022) developed a nominal land valuation model using the Best-Worst Method (BWM) and GIS. By weighting spatial and planning factors such as distance to the sea, transportation nodes, and noise levels, the authors created a land value map independent of market transactions (Ayalke & Sisman, 2022). This approach demonstrates the feasibility of using geospatial data for valuation in areas with limited sales data.

Similarly, Droj et al. (2024) proposed a valuation methodology based on mathematical appraisal models embedded in GIS platforms. Their study emphasized the role of layers such as land use, road access, and infrastructure availability in determining land value, enabling scalable and transparent appraisal systems(Droj et al., 2024). These spatially-integrated methods align with the approach in this thesis of incorporating adjacent land use/zoning data into valuation.

In these studies, Multi-Criteria Decision Analysis (MCDA) frameworks, such as AHP and BWM, are frequently employed to structure and weight the influence of each factor. While the study described in this thesis diverges in relying more on statistical and regression-based modeling, these literature insights validate the significance of structured spatial integration in developing fair and location-sensitive valuation tools.

2.2.2 Economic and Financial Modeling Approaches for Estimating Land Value

The second group of studies anchors land valuation in economic theory and statistical modeling. These works utilize hedonic pricing, multiple regression analysis (MRA), geographically weighted regression (GWR), and system dynamics to examine how different attributes, which include land use type, zoning classification, and neighborhood effects—influence parcel prices.

MRA serves as one of the most widely adopted empirical techniques in land valuation research that enables the estimation of the marginal contribution of multiple independent variables to observed land price while holding other factors constant. As demonstrated by Demetrious (2016), regression-based land valuation models provide a structured and transparent framework for decomposing parcel price into measurable spatial, regulatory, and locational components. Demetrious emphasizes that properly specified regression models enhance objectivity in land assessment by replacing purely expert-driven judgment with statistically interpretable relationships (Demetriou, 2016). Thus, the Hedonic regression pricing model deconstructs land prices into constituent attributes such as size, location, zoning, and permitted use, which allows for precise estimation of individual variable contributions. The method is especially useful in this thesis, where zoning, land use, and their adjacent features are treated as discrete value influences.

Similarly, the study "Land Use Regulation and Urban Land Value: Evidence from China" employed a Generalized Method of Moments (GMM) regression with panel data to estimate the effects of zoning changes on land prices across time and space. Their Two-Way Fixed Effects

model accounted for variations between cities and years, providing robust empirical support for the influence of policy-driven zoning shifts (Han et al., 2020).

Likewise, Greenaway-McGrey et al. (2021) applied hedonic regression to examine the premium effects of specific zoning and amenity features on property value, emphasizing that land use regulations can create both scarcity and benefit effects that are capitalized into prices (Greenaway-McGrevy et al., 2021). Another example, Wu et al. (2022), used MRA and GWR to explore the spatial heterogeneity in the relationship between zoning laws and land prices (Wu et al., 2022).

Together, these approaches highlight how statistical and spatial tools can be harmonized to estimate land value changes in response to planning interventions. In the case of Aurora, the integration of zoning density, adjacent land use types, and OP policy designations into a hedonic or MRA framework aligns well with these precedents.

2.3 Key Theoretical Questions and Scholarly Debates

While considerable progress has been made in modeling how planning policies affect land value, important theoretical questions and debates persist within the literature. Specifically, prior research has advanced from descriptive and case-based assessments toward formalized empirical models that quantify the capitalization of zoning regulations, land use controls, and policy interventions into land prices using hedonic regression, panel data methods, and spatial econometric techniques. Scholars continue to grapple with how zoning, land use regulations, and policy shifts are capitalized into land prices over time, and whether these interventions

consistently generate the intended outcomes in local markets. Additionally, divergent empirical findings reflect persistent methodological challenges, including data limitations, endogeneity, and the spatial heterogeneity of land markets. The following subsections examine these core questions and highlight areas of scholarly disagreement relevant to the design of this study.

2.3.1 Core Questions on Planning Interventions and Land Markets

A central line of inquiry in land valuation research concerns whether and how planning interventions, including zoning and land use regulations, are systematically reflected in land prices. As Kok, Monkkonen, and Quigley (2014) observe, regulations can operate through multiple channels, simultaneously restricting supply, conferring amenities, and signaling development expectations (Kok et al., 2014). This complexity raises foundational questions: Do planning controls inherently create scarcity premiums, or do they generate value primarily by reducing uncertainty and clarifying permissible development? Jalali et. al. (2022) further argue that regulatory measures can produce dynamic feedback loops over time, whereby constraints imposed in one period alter investment incentives and market perceptions in subsequent periods(Jalali et al., 2022).

Another core question concerns the scale at which these effects manifest. While some studies treat land use and zoning designations as parcel-level variables with direct price effects (Gore & Nicholson, 1991), others suggest that policy impacts are better understood in aggregate, through their influence on broader neighborhood and municipal land markets (Wu et al., 2022). The distinction between localized and systemic effects remains an important theoretical

consideration for valuation models, especially in suburban contexts like Aurora, where planning policies may have heterogeneous impacts across different development zones.

2.3.2 Conflicting Findings and Methodological Challenges

The literature also reflects considerable debate regarding the consistency and magnitude of policy effects on land values. For example, in their study of the San Francisco Bay Area, Kok et al. (2014) report that stricter land use regulations are associated with higher land values, attributing this to scarcity and signaling effects (Kok et al., 2014). In contrast, Freemark (2019) finds that upzoning Chicago increased property values but did not lead to measurable increases in housing supply over a five-year horizon, suggesting that policy interventions can produce short-term price capitalization without achieving broader affordability goals. Similarly, Rathnayaka and Prasada (2020) highlight that archaeological regulations had weaker and less consistent effects on land prices compared to zoning designations, pointing to the importance of regulatory clarity and enforceability (Rathnayaka & Prasada, 2020).

Methodologically, these divergent findings often reflect challenges in isolating causal effects. Studies employing hedonic regression may struggle with omitted variable bias and spatial autocorrelation (Kendall & Tulip, 2018; Rathnayaka & Prasada, 2020), while panel data approaches using fixed effects and instrumental variables (Freemark, 2020) introduce their own assumptions and limitations. Ongoing debates in the research community underscore the need for carefully specified models that account for both spatial heterogeneity and the dynamic nature of policy impacts on land markets.

2.4 Identified Research Gap: Parcel Level Regulatory Analysis in Outlying Communities such as Aurora within Large Urban Areas such as the GTA (Greater Toronto Area)

The preceding sections demonstrate that planning policies, including both OPs and ZBLs, are playing crucial role in shaping land values through mechanisms such as development control, density allowances, and regulatory signaling. Numerous studies have examined how these instruments function across global and national contexts. However, despite this growing body of literature, a critical gap remains in how regulatory effects on land value are experientially identified, interpreted, and different at the parcel level.

A key unresolved issue concerns the disconnect between price capitalization and policy outcomes. Empirical evidence shows that regulatory changes can be rapidly capitalized into property prices without producing corresponding changes in development or housing supply. For example, Freemark (2020) finds that upzoning in Chicago increased property values but did not lead to measurable increases in housing supply over a five-year horizon. This result highlights that price responses alone may not reflect the functional effectiveness of planning interventions, raising questions about how valuation models should interpret regulatory capitalization.(Freemark, 2020)

Relatedly, the literature reports inconsistent valuation responses across different types of regulations. Rathnayaka and Prasada (2020) demonstrate that archaeological and heritage regulations exhibit weaker and less consistent impacts on land prices compared to zoning designations, underscoring the importance of regulatory clarity, enforceability, and

development relevance. These findings suggest that not all planning policies generate equivalent valuation signals, and that models which treat regulatory instruments as uniform predictors risk oversimplifying market behavior(Rathnayaka & Prasada, 2020).

Taken together, these studies reveal a broader methodological gap: existing land valuation models often identify price effects without adequately distinguishing which regulatory mechanisms matter, under what conditions, and for which types of parcels. Limited attention has been given to separating land-driven valuation effects from housing-structure-driven price dynamics, or to integrating spatial regulatory context in a way that allows policy signals to be interpreted rather than merely detected. Addressing this gap requires valuation frameworks that explicitly link regulatory intent, development capacity, and parcel-level market response.

2.4.1 Limitations of Current Literature

Despite extensive research examining the relationship between planning regulations and land value, existing studies exhibit important limitations that can be broadly categorized into data-related constraints and modeling-related constraints.

A primary limitation of prior research lies in the availability and resolution of data used to evaluate regulatory impacts. Many studies rely on aggregated market indicators, assessed land values, or limited samples of transaction data, which restrict the ability to observe parcel-level valuation responses to planning policies. In some cases, regulatory effects are inferred indirectly due to the absence of detailed, longitudinal real estate transaction records linked to planning designations. As a result, empirical analyses often lack the granularity required to isolate how

specific regulatory changes are capitalized into individual parcel prices. The limited integration of comprehensive transaction datasets with parcel-level planning information constrains both the precision and interpretability of valuation outcomes.

In addition to data constraints, existing valuation models face methodological limitations in representing complex regulatory environments. Many studies examine zoning or land use policy in isolation, frequently operationalizing regulation through generalized land use categories without distinguishing between strategic planning instruments, such as OP, and site-specific ZBL. This simplification overlooks the hierarchical and interactive nature of planning frameworks, where zoning must conform to, yet may diverge from, higher-level policy intentions. Furthermore, most models incorporate a limited set of regulatory and spatial variables, often excluding adjacency conditions, parcel size heterogeneity, or the distinction between land-driven and structure-driven valuation regimes. As a result, existing models may identify price capitalization effects without adequately explaining the mechanisms through which planning policies influence land value.

Together, these data and modeling limitations highlight the need for valuation frameworks that combine high-resolution transaction data with integrated regulatory and spatial modeling approaches. Addressing these limitations is essential for improving the empirical evaluation of planning-policy impacts on land markets.

2.4.2 Justification for Localized, Spatial-Policy-Integrated Valuation

Given these limitations, there is a strong rationale for conducting a parcel-level analysis of planning policy impacts on land valuation within Aurora. Such an approach is justified on both conceptual and methodological grounds. Conceptually, it addresses the prevailing issue that most valuation models oversimplify regulatory environments, either by overlooking Official Plan guidance or by ignoring the practical enforcement and flexibility embedded in zoning bylaws.

Methodologically, the thesis responds to the absence of data-driven, factorized valuation models in the local planning discourse. Existing land value assessments in Aurora and the GTA tend to rely on appraisal expertise or planning rationale, rather than structured models that integrate regulatory factors with spatial and market data. As such, this research proposes a new empirical framework that uses categorized, interpretable planning variables to explain parcel-level valuation shifts.

In doing so, this thesis contributes to filling a significant empirical and methodological gap in literature by offering a localized, spatially nuanced, and policy-integrated land valuation model. This approach not only improves valuation precision in complex regulatory environments but also enhances the applicability of planning policy analysis in municipalities undergoing regulatory transformation and growth pressure.

2.5 Concept Models: Regression and Prediction Modeling

The analysis of land value and the evaluation of planning-policy impacts have long relied on quantitative modeling frameworks that relate observed market prices to underlying physical,

spatial, and regulatory attributes. Within the literature reviewed for this study, two methodological traditions are particularly relevant. The first is regression-based hedonic modeling, which provides an interpretable structure for estimating the marginal effects of land characteristics and planning variables on price. The second consists of prediction-oriented modeling approaches, which emphasize explanatory power and pattern recognition and are increasingly used to complement traditional econometric analysis. Together, these frameworks provide the conceptual foundation for examining how planning policies influence parcel-level land values.

2.5.1 Multi-Regression

Multi-regression-based hedonic price models remain the dominant analytical framework in land valuation research and form the methodological backbone of studies examining how land-use regulations are capitalized into market prices. The core assumption of hedonic modeling is that land price reflects the implicit value of a bundle of attributes, including parcel size, location, accessibility, surrounding land uses, and regulatory constraints. By decomposing price into these components, regression models allow researchers to estimate the marginal contribution of individual attributes while holding others constant(Demetriou, 2016).

Literature consistently highlights parcel size as a nonlinear determinant of land value. Ecker and Isakson (2005) demonstrate that for small residential parcels, price behavior is dominated by structural characteristics rather than land attributes, resulting in a change point at relatively small lot sizes(Ecker & Isakson, 2005). Their findings suggest that land value per unit area

declines sharply beyond this threshold, reinforcing the importance of nonlinear specifications and segmentation in regression models. Clauretie and Li (2017) extend this insight to development-scale parcels, showing that price-per-acre relationships vary substantially across larger parcel sizes and that optimal value occurs at a higher scale in less constrained urban contexts(Clauretie & Li, 2019). Together, these studies illustrate that linear regression models without size transformation or data set segmentation risk conflating house-driven and land-driven valuation dynamics.

Regression models are also widely applied to examine the effects of planning regulations and zoning controls. The literature reviewed in this study indicates that zoning designations, density allowances, and land-use restrictions are frequently significant predictors of land value, particularly for parcels with redevelopment potential (Clauretie & Li, 2019; Demetriou, 2016; Ecker & Isakson, 2005; Freemark, 2020). However, their explanatory power is highly context dependent and varies with parcel size and market structure. This reinforces the need for regression frameworks that incorporate interaction terms, polynomial transformations, or segmented specifications to accurately capture regulatory effects.

A key strength of regression-based approaches lies in their interpretability. Coefficient estimates provide direct insight into the direction and magnitude of relationships between planning variables and land value, making regression models particularly suitable for policy analysis and planning applications. For this reason, regression modeling remains central to studies seeking to evaluate planning-policy impacts in a transparent and replicable manner.

2.5.2 Prediction Modeling

In contrast to multi-regression-based methods, prediction-oriented modeling (e.g. trained AI networks) approaches emphasize cross-sectional price prediction and pattern recognition, rather than coefficient interpretation. In this study, prediction refers to the estimation of parcel-level value outcomes given observed spatial, regulatory, and market attributes at a specific point in time, rather than predicting future price across time. While the literature reviewed in this study does not position predictive models as replacements for regression frameworks, it increasingly recognizes their value as complementary analytical tools. These approaches are particularly effective in capturing nonlinear relationships and complex interactions that may be difficult to specify explicitly in traditional econometric models.

Within the land-valuation literature, prediction models are commonly used to test the robustness of variable selection and to assess whether observed relationships persist when functional-form assumptions are relaxed. Studies comparing regression and predictive approaches suggest that although predictive models may achieve higher explanatory performance, they often sacrifice interpretability, limiting their direct applicability for policy evaluation (Clauretie and Li, 2017). As a result, prediction-oriented models are most effective when used alongside regression analysis, serving as validation tools rather than standalone policy instruments.

The literature further emphasizes that predictive models are sensitive to data composition and market structure. When datasets are dominated by a particular parcel type, such as small residential lots, predictive algorithms tend to reflect the dominant valuation pattern and may

underrepresent regulatory effects that occur primarily in development-scale parcels (Ecker and Isakson, 2005). This insight highlights the importance of combining prediction models with data segmentation strategies to ensure that planning-policy effects are not obscured by structural market characteristics.

Overall, the literature supports a combined analytical framework in which regression models provide interpretive clarity and policy relevance, while prediction-oriented models enhance robustness and confirm the stability of observed relationships. This complementary perspective directly informs the methodological design of this study and supports the integration of both regression-based and predictive approaches in evaluating planning-policy impacts on land value.

Chapter 3 Proposed Methodology

This chapter outlines the methodological framework developed to examine how planning policies, physical and spatial characteristics influence parcel-level land values in the Town of Aurora (See Figure 5). Building upon the theoretical and practical foundations discussed in Chapter 2, the methodology integrates multi-regression-based hedonic modeling with contemporary machine-learning techniques to provide both interpretive and predictive insights into the determinants of land value.

The chapter begins by describing the data sources and preparation processes used to construct the analytical dataset. It then presents the selection and definition of variables, highlighting how each represents a unique dimension of the land-valuation process, including physical parcel attributes, policy designations, adjacent conditions, and accessibility measures. The subsections introduce the modeling framework, beginning with the specification of regression models used to quantify the relationships between land value and planning variables. This is followed by an explanation of the complementary machine-learning model, LightGBM, which is applied to validate and enhance interpretive findings through predictive analysis.

By combining statistical inference with machine-learning validation, this chapter provides a rigorous and replicable framework for estimate land value within a policy and planning context. The results derived from this methodology form the basis for the analytical discussion presented in Chapter 4.

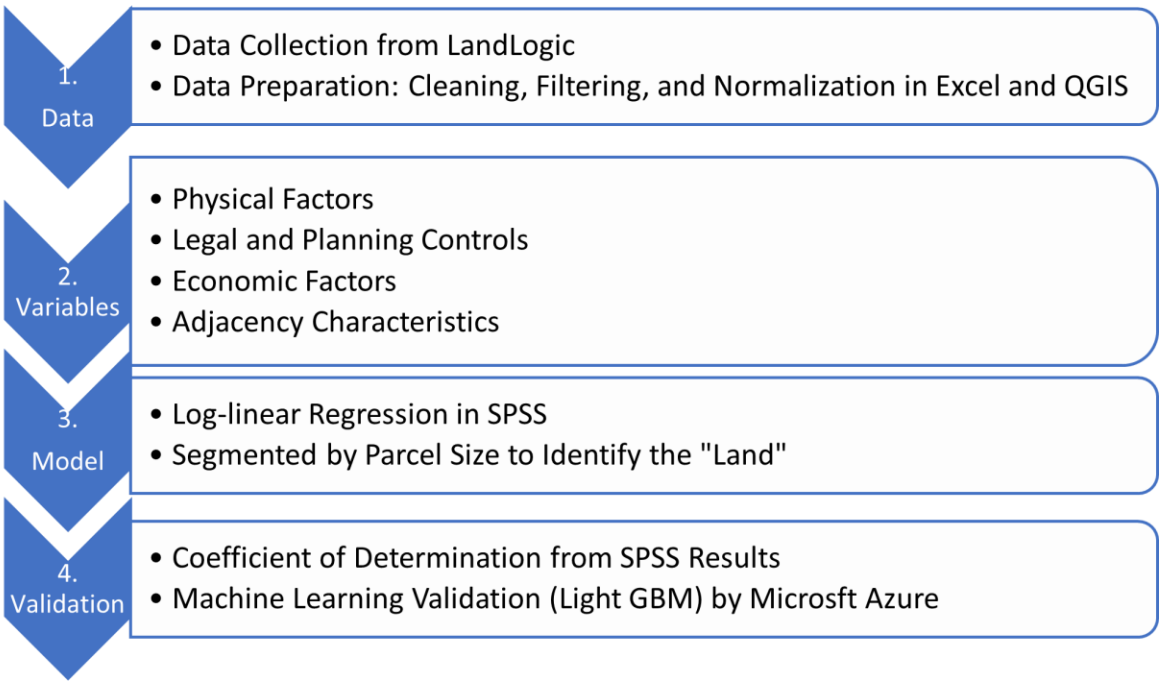


Figure 5 Overall Research Variables and Methodological Framework

3.1 Data Collection and Preparation

In this section, the various data sets used are described, including their sources and the processes of variable generation through spatial aggregation and calculation. Note that this dataset only contains parcel-level planning and spatial variables but does not include detailed structure-level attributes such as dwelling floor area or building condition. Consequently, for parcels containing existing residential structures, observed transaction prices represent a combination of land and building value rather than pure land value.

The dependent variable in this study is derived from parcel-level real estate transaction data for the Town of Aurora covering the period 2014 to 2025. The testing data originates from LandLogicTM proprietary transaction database (LandLogic, 2025), which integrates MLS (multiple Listing Service) records with geospatial parcel information. The raw data includes attributes of MLS Number, address, sold price, sold data, listing entry date, parcel geometry, longitude, and latitude (Table 2). These records capture sales across a full range of residential, commercial, and industrial properties.

Since MLS transactions generally reflect the combined market value of both land and built structures, distinguishing the proportion attributable solely to land presents an analytical challenge. Following expert guidance (Amin, 2025), parcels were classified according to size to separate primarily land-driven transactions from structure-dominated ones. Parcels with an area greater than or equal to 0.3 acre were treated as land-dominant values, while those smaller than 0.3 acre were considered property-dominant and therefore excluded from the land-value analysis. While the principle behind this division is clear and generalizable as part of

the research methodology applied, the exact value assigned depends on the region being modeled and its typical lot sizes. Thus, some expert judgment is required at this point. Alternatively, brute force sensitivity analysis could be used to find an optimum value.

In total, 11,636 real estate transactions were available for the 2014–2025 period. To focus the analysis on policy-relevant sites, the dataset was restricted to 8,298 parcels whose OP land-use designation changed between the 2010 and 2024 OP documents. The ideal comparison would involve evaluating land prices before and after the planning designation change. However, in practice, purchasers often anticipate policy shifts years before the formal adoption of the OP amendment. As a result, the "before-and-after" price comparison may span too long a period to isolate the true policy effect. In addition, an ideal longitudinal time-series dataset would further improve the accuracy of policy impact evaluation. However, historical planning-policy records and OP maps in Aurora are not consistently available in digital format. Even the 2010 and 2024 OP land-use maps required substantial digitization and processing efforts by LandLogic™ team. As a result, constructing a consistent time-series planning dataset for Aurora was not feasible within the scope of this study. To address this limitation, the study focuses on parcels whose land use designations changed between 2010 and 2024 OP based on data provided by LandLogic, in order to identify the general pattern of how planning policy adjustments influence land price. In addition, by connecting the changed land use parcels and the transaction data, the "Join attributes by locations" tool was used in QGIS. Furthermore, from this subset, a random sample of 300 parcels was drawn by assigning random variables using formular formula *= rand()* in Excel and selecting the lowest 300 records. After removing duplicate transactions at the same address (e.g., apartment-unit sales), the final analytical sample consisted of 278 transactions.

Expanding the analytical sample beyond 300 parcels was constrained by the intensive manual progressing required for several independent variables. Variables such as adjacency characteristics, zoning designation matching, and OP land-use identification required substantial manual verification and spatial processing because QGIS was unable to efficiently process larger datasets within the limited study timeline. For example, linking zoning designations and OP land-use information to parcel transactions records required approximately one week of processing. As a result, the selected sample size represents a practical balance between analytical depth, data quality, and project feasibility.

To ensure comparability and temporal consistency, all sale prices were converted to Canadian dollars per acre and normalized to 2025 price levels using the MLS Home Price Index (HPI) (Canadian Real Estate Association, 2025). This standardized the dependent variable by inflation-adjusted land price per acre and serves as the outcome measure in all subsequent regression models.

Table 2 LandLogic™ Source Data Attribute

Sources	Variable	Full Name	Definition	Relevance to Land Valuation
2014-2024 Real Estate Transaction Records	LOCATION	Location	Address of the parcel	Helps to locate the parcel
Official Plan for Town of Aurora, September 2010	Lu_2010_all	Land Use in 2010	Land Use designation of the parcel according to the Aurora 2010 Official Plan	Helps to identify the land use designation change from 2010 to 2024.
Official Plan for Town of Aurora, September 2024	Lu_2024_all	Land Use in 2024	Land Use designation of the parcel according to the Aurora 2024 Official Plan	Helps to identify the land use designation change from 2010 to 2024.

LandLogic™ proprietary transaction database	MlsNumber	MLS Number	A unique identifier for each property on housing market (marketer et al., n.d.).	Helps to identify each transaction and ensures spatial accuracy through linkage with the QGIS layer.
	SoldPrice	Sold Price	The real estate transaction price.	Helps to identify the physical structure and land value of the property.
	SoldDate	Sold Date	The real estate transaction closing date.	Helps to identify the transection closing date.
	Area_Acres	Area in Acres	Denotes parcel size, measured in acres.	Helps to identify the land size in a unified unit.

3.2 Variable Selection

This section introduces the variables used in the regression models and explains their theoretical and empirical demonstration. The variables were selected based on their relevance to land value formation in Aurora and their consistency with findings in previous planning economic studies. Each variable represents a distinct dimension of land valuation, encompassing parcel characteristics, planning policy factors, spatial context, and accessibility. The preparation of these variables was completed in Excel, and subsequent analyses were performed in SPSS.

3.2.1 Data Preparation and Variable Selection

This study uses a combination of continuous, categorical, and dummy variables to capture the multifaceted influences on land value. The dependent variable (See Table 3) represents the

inflation-adjusted land transaction price per acre, while the explanatory variable reflects the physical, policy, and location attribute of each parcel.

The dependent variable was derived from LandLogic™ parcel-level transaction dataset for Aurora (2014–2025) and normalized to 2025 price levels using the HPI (Canadian Real Estate Association, 2025). The log-transformed normalized price per acre serves as the primary outcome variable in all regression analyses.

Table 3 Dependent Variables

Sources	Variable	Full Name	Definition	Relevance to Land Valuation
Computed as SoldPrice divided by Area_Acres	Price per Acre	Sold Price in Canadian Dollars per Acre	The unified sold price in Canadian dollars per Acre.	Helps to bring the land value into the same unit to compare.
Computed from the Year in SoldDate	Sold Year	Sold Year	The real estate transaction is closing year.	Helps with the calculation of normalized land price.
Computed from the Price per Acre by Using HPI and Sold Year.	Normalized Price (2025)	Normalized Sold Price in Canadian Dollars per Acre in 2025 level	The sale price adjusted to 2025 Canadian dollars using the HPI.	Helps to bring the land value into the same year to compare, which avoids the housing price changes in different years.
Computed from Log-transformed Normalized Price (2025)	Normalized Price per Acre (Log10)	Log10 Transformed Normalized Sold Price in Canadian Dollars per Acre in 2025 level	The log-transformed dependent variable is used in all regression models.	Helps to reduce the extreme outliers impact on the model.

Dependent variables were identified through both theoretical guidance and empirical testing. Preliminary regressions revealed notable nonlinearity in the relationship between parcel size and price, especially among larger parcels where predicted values deviated substantially from

observed prices. To address this issue, the parcel size variable (Area in acres) was scaled by dividing each value by the largest parcel size in the dataset (37.026 acres). Three polynomial terms, which are Area, Area^2, and Area^3, were then generated to capture the nonlinear effects of size. This approach aligns with evidence from prior research that emphasizes scale-dependent land-value behavior. Ecker and Isakson (2005) observed a threshold around 0.08 acres for residential-scale parcels, while Clauretie and Li (2017) identified a turning point of approximately 14 acres for large development parcels. In Aurora, the average parcel size is 0.48 acres, and the median is 0.117 acres, indicating a distinct context compared with these studies. In addition, guided by expert judgment (Amin, 2025) and empirical size distribution analysis, which showed that 89 percent of parcels are smaller than 0.3 acre, this study adopted 0.3 acre as a threshold to separate house-driven and land-driven value. Two segmented regression models were developed accordingly.

The independent variables are organized into four categories based on Bencure et al.'s structure: (1) physical factors, (2) legal and planning controls, (3) economic factors, and (4) adjacent characteristics. Their definitions, units, and expected relationships with the dependent variable are summarized in Table 4.

Table 4 Independent Variables

Category	Variable	Full Name	Definition	Relevance to Land Valuation
Physical Factors	Size (Small = 1, Medium = 2, Large = 3)	Size Category of the Parcel	Classifies parcels by area (<1 acre, 1–5 acres, >5 acres).	Based on expert's experience, that general classified the land size's impact on its value. It only applies to parcels bigger than 0.3 acres

				to detailed size impact on land.
Legal and Planning Controls	Residential (1 = Yes, 0 = No)	Residential Land Use Designation Category	Land use designation in 2024 Aurora OP includes Suburban, Estate, Medium-High Urban, Low-Medium Urban, Stable Neighbourhoods, and Mixed-Use areas.	The land use designation from Aurora's OP; helps to identify the land use designation impact on the land price.
	Commercial (1 = Yes, 0 = No)	Commercial Land Use Designation Category	Land use designation in 2024 Aurora OP includes Community Commercial, Service Commercial, Convenience Commercial, and Major Retail Centres	
	Employment (1 = Yes, 0 = No)	Employment Land Use Designation Category	Land use designation in 2024 Aurora OP includes Industrial Lands and Business Park	
	Restrictive Use (1 = Yes, 0 = No)	Restrictive Land Use Designation Category	Land use designation in 2024 Aurora OP includes Environmental Protection area and Major Transit Station Area (MTSA). Environmental covers Environmental Protection, Oak Ridges Moraine Natural Linkage, Countryside, Natural Core, and Parkland areas. MTSA Boundary indicates parcels within or adjacent to Aurora's Promenade and other MTSA Mixed Use zones.	

	Service (1 = Yes, 0 = No)	Service Land Use Designation Category	Land use designation in 2024 OP includes Community Service and Facilities land uses.	
	Percentage of Useable Land	Percentage of Developable Land without any Additional Permit Needed in the Parcel	The Proportion of land available for developer without additional conservation approvals.	Additional conservation approval requirement within the parcel increase the development time cost, which reduce the land price.
	Density	The maximum density of the parcel designed by the zoning by-law.	The estimated Floor Area (FAR) derived from the applicable zoning by-law.	Higher maximum density increase the land price.
Economic Factors	Access to Road (Level 0-9)	Level of the Road Can Access to the Parcel	Represents the highest class of road accessible from the parcel, with codes ranging from 0 (no access) to 9 (freeway).	Access to road will reduce the development time and additional permit required of the parcel, which increases the land value.
	Distance to City Core (km)	The Straight-line Distance from the Parcel to Aurora's Downtown Core	Measures the straight-line distance from the parcel centroid to the intersection of Yonge Street and Wellington Street, representing Aurora's downtown center.	The closer to the city core, the higher land value is.
Adjacency Characteristics	Sum of Adjacent Lands	Sum of the Number of Adjacent Lands of the Parcel	The number of parcels directly abutting the target parcel.	As the target parcel has more adjacent parcels, more stakeholders are involved in developing progress.
	Percentage of Adjacent Residential	Percentage of Adjacent Parcels Under Residential Category Land Uses	The percentage of residentials use parcel directly abutting the target parcel from all types of adjacent lands.	As the target parcel has more adjacent residential parcels, the target parcel is more likely located in a stable neighborhood which is less likely to change land use type.

Collectively, these variables form a comprehensive framework for quantifying the relationship between land-use policy, spatial context, and land value in Aurora. Their inclusion enables the two separate regression models to distinguish between intrinsic parcel characteristics and planning-driven influences, supporting a more robust evaluation of how OP changes have shaped local land valuation patterns.

3.2.1.1 Physical Factors

Physical characteristics describe the inherent attributes of each parcel that directly affect its market value and development potential. These factors are foundational to any land-valuation model because they represent the basic physical constraints and opportunities associated with the land itself (Bencure et al., 2019).

The most fundamental of these attributes is parcel size, which reflects both the scale of potential development and the degree to which value is influenced by existing structures. First, parcels were segmented into house-driven and land-driven valuation regimes using a threshold of 0.3 acre. Second, parcel size was categorized into three levels based on professional practice (Sarang, 2025) to further capture the differentiated impact of parcel size on land value. In addition, prior research (Ecker & Isakson, 2005; *Land Values: Size Matters | The Journal of Real Estate Finance and Economics*, n.d.) has shown that the relationship between size and price per acre is nonlinear, as smaller parcels often reflect housing-driven value while larger parcels

capture land-driven value. Therefore, parcel size in Scaled and its polynomial terms (Scaled, Scaled^2, Scaled^3) are included to account for this nonlinearity.

Thus, these physical factors represent the inherent characteristics that determine the parcel's value independent of regulatory or locational influences.

3.2.1.2 Legal and Planning Controls

Legal and planning controls are central to this study because they capture how municipal policy frameworks and zoning designations change influence market expectations and land value. These variables translate planning theory into measurable, spatially explicit factors that can be analyzed quantitatively.

The key variables in this group include OP designations and density representing the zoning by-law. These designations define the permissible uses, density limits, and long-term growth intent of the municipality. Previous studies (Demetriou, 2016; Kendall & Tulip, 2018) have shown that such regulatory instruments can substantially affect land value by altering development potential and perceived risk.

Dummy variables are therefore created for Residential, Commercial, Employment, Business Park, Service, and Restrictive designations. Including these policy-based variables allows the model to quantify how land value responds to zoning and OP amendments. For example, if the parcel is within the Suburban land use designation, it has been marked as "1" in dummy variable. Thus, this focus on parcels that experienced land-use designation changes between 2010 and 2024 Official Plans further isolates the planning-policy changes' effect on land value.

In addition, consolidating detailed land-use designations into broader land-use categories reduces the number of dummy variables required in the model, while the inclusion of the density variable provides finer differentiation among land uses within these generalized categories.

Other legal and planning controlled variables, including density and percentage of usable land, are capturing development intensity from the ZBL and physical constraints from restriction policies. Density (measured as Floor Area Ratio, FAR) indicates the scale of permissible construction, which directly contributes to potential land yield and thus to higher value. However, it is not explicitly specified in the ZBL. It was calculated through multiplying the maximum permitted lot coverage by the estimated number of stories, derived from the maximum building height. For example, in the C1 Convenience Commercial zone, the maximum lot coverage is 30 percent, and the maximum height is 10 m. Assuming an average commercial story height of 3.3 m, the estimated number of stories is three, resulting in an estimated FAR of approximately 1.05. Besides, the percentage of usable land measures the proportion of the parcel that can be developed without additional approvals from conservation or environmental authorities, reflecting practical limitations on the parcel's development capacity. It was estimated manually using LandLogic™, by overlaying environmental restriction layers (e.g. Environmental Conservation Authority Restricted Area) on top of the changed parcels.

This group of variables operationalizes the regulatory dimension of value formation, linking planning instruments to market outcomes in a replicable way for other municipalities with accessible OP and zoning datasets.

3.2.1.3 Economic Factors

Economic factors represent the market-driven influences that capture demand, accessibility, and broader locational dynamics affecting land value. In land-valuation models, these variables often serve as proxies for the parcel's economic connectivity and potential profitability.

The most direct economic variable in this study is distance to the city core, which approximates access to central amenities, employment, and infrastructure. Urban economic theory (Alonso, 1964; Fujita, 1989) suggests that land value generally declines with distance from the central business district, as accessibility costs increase. In Aurora, the intersection of Yonge Street and Wellington Street is identified as the central reference point used in LandLogic platform for this distance measure.

The access-to-road variable further reflects transportation connectivity, capturing the level of integration within the regional mobility network. It ranked from 0 to 8 manually in LandLogic platform, from small to large including No Access, Laneway, Private Road, Proposed Road, Rural Road, Urban Road, Regional Road, and Arterial Road. In other words, parcels located on arterial or regional roads typically command higher values due to greater visibility and accessibility, whereas those accessible only by private or laneway roads may experience reduced market exposure.

Additionally, density and adjacency to commercial or employment lands indirectly represent economic potential by indicating proximity to active market areas. These economic variables

together form the spatial-economic layer of the valuation model, which can be easily replicated in other municipalities by substituting local road and core-distance datasets.

3.2.1.4 Adjacency Characteristics

Adjacent characteristics describe the spatial context of each parcel and its immediate surroundings. These variables capture how nearby land is used and neighborhood composition influence parcel desirability and perceived market potential.

Following experts' experience (Amin, 2025; Saeid, 2025) and planning documents (Appraisal Reports) in urban and environmental economics, adjacency effects are modeled as percentage indicators, representing the percentage of neighboring parcels categorized as residential, commercial, industrial, restrictive, business park, or service uses. These percentages, using dummy variables 0 and 1, are calculated manually based on counting parcels sharing a common boundary with the target parcel by using LandLogic.ai platform and divided by the sum of adjacent parcels.

The rationale for including adjacency variables is twofold. First, land values are rarely determined in isolation, that they reflect the combined attributes of both the parcel and its context. For instance, adjacency to residential or commercial uses may enhance value by signaling compatibility and development demand, while adjacency to industrial or restrictive uses may reduce market appeal. Second, adjacency metrics are highly transferable: other municipalities can replicate this approach using their parcel and zoning layers, applying the same proportional calculation method.

By incorporating adjacency characteristics, the model integrates micro-spatial context into land valuation, ensuring that the surrounding urban fabric is considered alongside intrinsic and policy-related parcel attributes.

3.2.2 Model Specification

A wide range of quantitative models have been applied in land-valuation research to estimate the relationship between parcel characteristics and market price. As reviewed in Section 2.5, these approaches can generally be grouped into two methodological streams: traditional econometric regression models and contemporary data-driven predication models. Each has both advantages and limitations when applied to spatial land-value analysis.

Classical regression models such as hedonic price models, have long served as the foundation for land-value estimation in urban economics and real estate studies. The hedonic approach divided a property's value into the implicit value of its measurable attributes, like parcel size, location, zoning, and accessibility (Ai et al., 2020; Bencure et al., 2019; Demetriou, 2016; Jafary et al., 2024). Numerous studies, including Ecker and Isakson (2005), Demetrious (2016), and Clauretie and Li (2017), have demonstrated that regression-based models are effective for identifying statistically significant factors that influence land or property values. These models are also suitable for testing hypotheses related to planning and policy effects, as they provide interpretable coefficients and allow for statistical inference regarding the direction and magnitude of variable impacts (Clauretie & Li, 2019; Demetriou, 2016; Ecker & Isakson, 2005).

3.3 Regression Modeling Approach

This section outlines the analytical framework used to examine the relationship between land price and planning, spatial, and physical characteristics of parcels in Aurora. The regression model is designed to quantify the magnitude and direction of the influence of OP changes, zoning, adjacency, and accessibility factors on normalized land values. The modeling approach follows the hedonic valuation tradition in urban economics while adapting to the unique characteristics of the dataset compiled for this study.

3.3.1 Model Structure and Rationale

The MLR model was selected as the primary analytical technique. The MLR framework (Equation 3.3.1.1) enables decomposition of parcel price per acre into the implicit value contributions of its attributes and allows for the assessment of policy-driven effects through statistically interpretable coefficients. Due to the dependent variable exhibits of a wide range of values across parcel types and years, the sale price per acre was log-transformed to reduce skewness and stabilize variance (Equation 3.3.1.2). This transformation also ensures that coefficient estimates represent percentage changes in land value associated with one-unit changes in explanatory variables.

$$P_i = \alpha + \sum_{k=1}^{K} \beta_k X_{ki} + \varepsilon_i \quad 3.3.1.1$$

where

P_i= observed price (or price per acre) of parcel *i*

α = intercept (constant term) of the regression model

X_{ki}= parcel characteristics (size, land use, adjacency, etc.)

β_k= estimated coefficients

ε_i= random error term

$$Log_{10}(PA_i) = \alpha + \sum_{k=1}^{K} \beta_k X_{ki} + \varepsilon_i \qquad 3.3.1.2$$

where $PA_i$is the normalized sale price per acre (2025 CAD), log-transformed to stabilize variance and reduce the influence of outliers.

The use of a log-linear specification is consistent with previous empirical studies of land and property valuation (Demetriou, 2016; Kendall & Tulip, 2018), which demonstrate that property values typically respond proportionally rather than linearly to incremental changes in explanatory factors.

3.3.2 Segmented Regression by Parcel Size

As discussed in Section 3.2.1, parcel size plays a critical role in differentiating markets where price is driven primarily by building characteristics versus underlying land potential. To account

for this heterogeneity, the analysis applies a segmented regression framework separating parcels into two size groups:

- Small parcels (< 0.3 acre), representing primarily house-driven transactions;
- Large parcels ($\geq$ 0.3 acre), representing land-driven transactions.

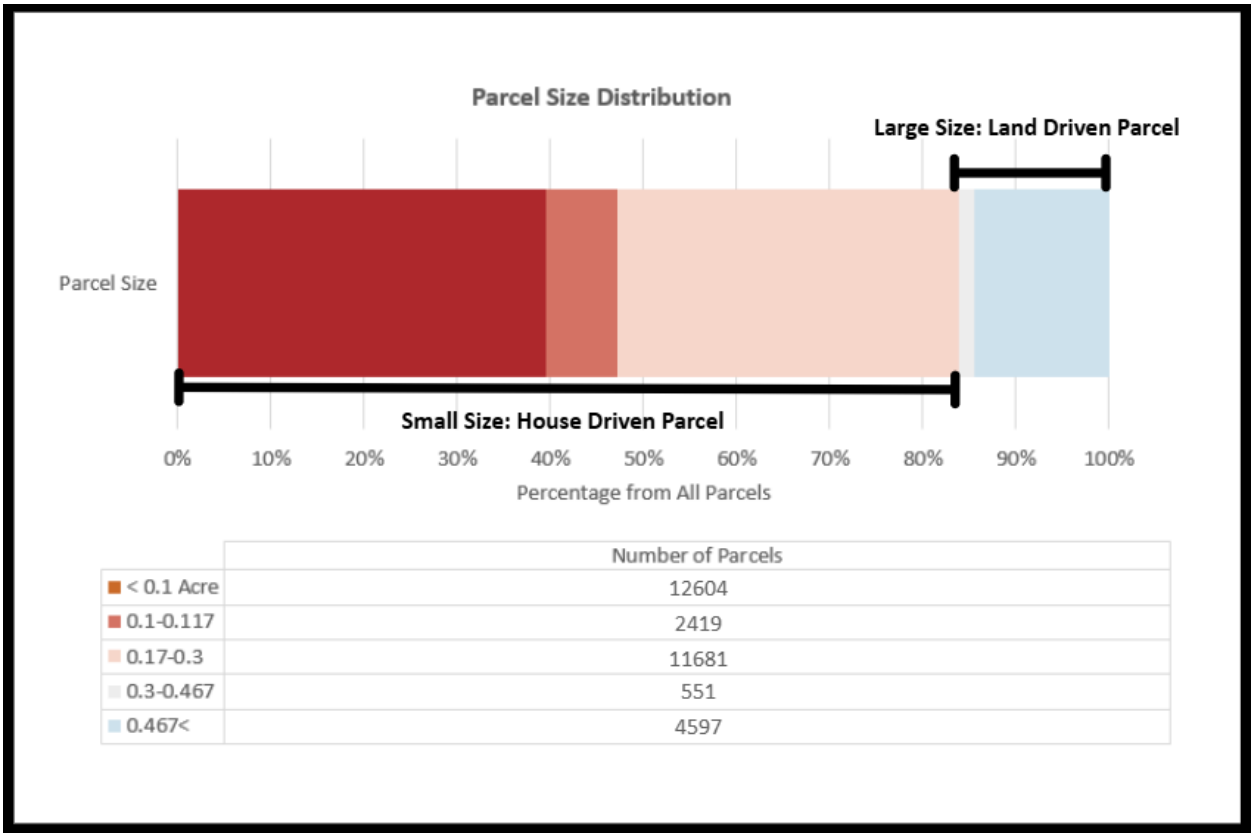
Separate regression models (Equation 3.3.2.1) were estimated for each segment to isolate the effects of planning variables within more homogeneous market contexts.

First, the empirical size distribution of Aurora's parcels, where the average and median parcel area is 0.467 and 0.117 acre, and approximately 89 percent are smaller than 0.3 acre (See Figure 6). Thus, the change point suggested in other study (Ecker & Isakson, 2005) does not fits to Aurora, the 0.3 acre threshold was determined based on expert insight (Amin, 2025) would be the best fit change point of the study. Moreover, this segmentation improves model interpretability and mitigates bias arising from nonlinear size effects observed in preliminary analyses.

Second, within each regression model, parcel size is introduced as a continuous variable through scaled polynomial transformations (Scaled, Scaled^2, and Scaled^3). This transformation improves numerical stability and reduces multicollinearity while allowing the model to capture nonlinear relationships between parcel area and land value.

Third, in the land-driven segment, an additional categorical Parcel Size variable (coded as small = 1, medium = 2, and large = 3) is included to further differentiate valuation patterns across development-scale parcels. This variable supports finer interpretation of how redevelopment

potential varies within the land-driven market and helps isolate transitional valuation behavior between medium and large development sites.



	Number of Parcels
< 0.1 Acre	12604
0.1-0.117	2419
0.17-0.3	11681
0.3-0.467	551
0.467<	4597

Figure 6 Parcel Size Distribution

$$\begin{cases} log_{10}(PA_i) = \alpha\hat{}(S) + \Sigma(\beta_k\hat{}(S)\, X_{ki}) + \varepsilon_i, \ if \ Area_i < 0.3 \\ log_{10}(PA_i) = \alpha\hat{}(L) + \Sigma(\beta_k\hat{}(L)\, X_{ki}) + \varepsilon_i, \ if \ Area_i \geq 0.3 \end{cases} \quad 3.3.2.1$$

Where:

$\log_{10}(PA_i)$= log-transformed normalized land price per acre for parcel *i*

$\alpha^{(S)}$, $\alpha^{(L)}$= intercepts for small (<0.3 acre) and large (≥0.3 acre) parcel groups

$\beta_k^{(S)}$, $\beta_k^{(L)}$= coefficients for explanatory variables X_{ki}in each segment

ε_i= random error term

3.3.3 Variable Transformation and Scaling

Continuous variables such as parcel area, adjacency proportions, and distances were standardized to ensure comparability and minimize scale-related estimation bias. Parcel area was further represented through scaled polynomial terms (Scaled, Scaled², Scaled³) to capture nonlinear relationships between land size and price. Dummy variables were applied for categorical designations including land-use types, zoning, and policy constraints. All spatial and policy variables were entered simultaneously into the regression to capture the combined influence of physical and regulatory factors on land value (Equation 3.3.3.1).

$$\begin{cases} log_{10}(PA_{\mathrm{i}}) = \alpha\text{^}(S) + \Sigma(\beta_{\mathrm{k}}\text{^}(S)\, X_{\mathrm{ki}}) + \varepsilon_{\mathrm{i}}, \ if\ Area_{\mathrm{i}} < 0.3 \\ log_{10}(PA_{\mathrm{i}}) = \alpha\text{^}(L) + \Sigma(\beta_{\mathrm{k}}\text{^}(L)\, X_{\mathrm{ki}}) + \varepsilon_{\mathrm{i}}, \ if\ Area_{\mathrm{i}} \geq 0.3 \end{cases} \qquad 3.3.3.1$$

Where:

$\log_{10}(PA_i)$= log-transformed normalized land price per acre for parcel *i*

$\alpha^{(S)}$, $\alpha^{(L)}$= intercepts for small and large parcels

$\beta_k^{(S)}$, $\beta_k^{(L)}$= coefficients for explanatory variables X_{ki}in each size segment

ε_i= random error term

3.3.4 Model Identification and Validation in Machine Leaning Model Light GBM

Standard identification procedures were performed to evaluate the reliability of the regression results. Model residuals were tested for normality and homoscedasticity to verify compliance with ordinary least-squares (OLS) assumptions. Influential outliers were identified through standardized residual and Cook's-distance analyses, and necessary adjustments were made to maintain model robustness.

To enhance interpretive confidence, the results from the regression models were cross validated using a Random Forest algorithm. While the regression provides coefficient-based interpretability, the Random Forest model serves as a non-parametric benchmark to evaluate variable importance and to verify whether the ranking of influential predictors aligns with the linear model's outcomes. Although Random Forest models are effective in capturing nonlinear relationships and performing robustly under imbalanced data conditions, they do not provide the same level or interpretive transparency as regression-based explanatory models, where the direction, magnitude, and statistical significance of planning-policy variables can be explicitly examined. Thus, dual-model strategy integrates statistical inference with machine-learning validation, strengthening both the explanatory and predictive aspects of the analysis.

In summary, the regression modeling approach combines traditional econometric estimation with modern predictive validation. The log-linear multiple regression identifies the direction and magnitude of the effects of land-use, zoning, and spatial variables on parcel-level land values,

while the segmented framework accounts for heterogeneity between small and large parcels. The complementary use of Random Forest analysis provides an additional diagnostic perspective on variable significance. The following chapter presents the regression results and interprets their implications for understanding how planning policies and spatial factors shape land valuation patterns in Aurora.

Chapter 4 Results

This chapter presents and interprets the results of the land-value modeling conducted for the Town of Aurora.

The analysis builds on the methodological framework introduced in Chapter 3, which is applying a segmented regression approach to examine how planning, spatial, and physical variables influence normalized land prices. The two regression models were estimated separately for parcels smaller than 0.3 acre (representing house-driven values) and for parcels equal to or larger than 0.3 acre (representing land-driven values). In addition to the regression analysis, a machine-learning model based on Light GBM was implemented to test the predictive accuracy and relative importance of the explanatory variables and to validate the robustness of the regression-based findings.

Model diagnostics, coefficients, and significance levels are discussed for each model, followed by a comparative interpretation of key findings. Results from the LightGBM model are subsequently integrated to assess consistency between statistical inference and predictive performance, thereby strengthening confidence in the identified relationships and supporting the overall analytical framework.

4.1 Segmented Regression Results

To evaluate how planning policies, spatial characteristics, and parcel attributes influence land values across different market conditions, two segmented regression models were estimated using the log-transformed normalized price per acre as the dependent variable. Segmentation was based on the 0.3-acre threshold identified in Chapter 3, with parcels smaller than 0.3 acre representing house-driven markets and parcels equal to or larger than 0.3 acre representing land-driven markets. The results for each model are presented below, followed by a comparison.

4.1.1 Model A: Parcels < 0.3 Acre (House-Driven Value)

Model Performance

The regression model estimated for parcels smaller than 0.3 acre demonstrates strong overall explanatory power. The model achieved an R value of 0.894, indicating a high degree of correlation between the observed and predicted values of normalized land price per acre. The adjusted R^2 value of 0.799 suggests that approximately 80 percent of the variation in the dependent variable is explained by the included predictors after accounting for model complexity. This level of explanatory performance indicates a stable model fit and suggests that overfitting is unlikely despite the inclusion of multiple spatial and policy-related variables.

Table 5, which is the Diagnostic evaluation of model residuals further supports the adequacy of the specification. Visual inspection of residual plots indicates approximate normality, with no systematic patterns that would suggest major violations of regression assumptions. While minor heteroscedasticity is present, its magnitude is limited and does not materially affect coefficient

interpretation. Thus, these diagnostics confirm that the regression framework is suitable for analyzing valuation behavior within the House-Driven Value category.

Table 5 Diagnostic Evaluation of Model Residual

	Number of Parcels	This Segmented Regression Model Outcomes	Expected Outcomes
within ±1σ	163	54.9%	68.0%
within ±2σ	296	99.7%	95.0%
within ±3σ	297	100.0%	99.7%

Coefficient Summary

Figure 7 presents the unstandardized coefficients, associated standard errors, and confidence intervals for Model A. Among the explanatory variables, distance to the city core emerges as the most statistically significant predictor. The coefficient for distance to the city core is positive and statistically significant at the 0.001 level, indicating that normalized price per acre increases with greater distance from Aurora's central area within the house-driven value parcel subset. This result reflects the spatial structure of Aurora's residential market, where newer suburban housing located farther from the historic core may command higher prices due to larger dwelling sizes, newer construction, and neighborhood amenities that are not directly captured in land-based variables.

Parcel size effects, represented through scaled polynomial terms, also contribute meaningfully to the model. Scaled parcel size refers to the normalized parcel-area variable used in the regression model to stabilize the polynomial specification. The Scaled parcel size is statistically significant and negative, suggesting diminishing marginal land value per acre as parcel size increases within the small-parcel range. Scaled parcel size's square and cube (Scale 2) are not

statistically significant, indicating that nonlinearity beyond the scaled parcel size effect is limited in this segment. This pattern is consistent with expectations for house-driven markets, where marginal changes in land area do not translate proportionally into price changes once a minimum lot size threshold is met. However, building-level characteristics such as house floor area were not included as a dependent variable due to the primary objective of this study, which is evaluating parcel-level land valuation and the influence of planning policies. Moreover, consistent and reliable building size data were not available for all transactions within the dataset. As a result, this model focuses on parcel attributes and regulator conditions to capture the mechanisms through which planning policies influence land value.

By contrast, planning-policy and regulatory variables show limited explanatory power in this model. Dummy variables representing residential designation, commercial and service designation, and restricted land designation are not statistically significant. Similarly, density, adjacency composition variables, accessibility to the road network, and the proportion of easily usable land do not exhibit statistically significant coefficients. These results suggest that, within the small-parcel market, regulatory and spatial planning variables are largely overshadowed by factors associated with the built structure, which are not explicitly modeled in this study.

Table 6 Coefficients from SPSS of Model A: Parcels < 0.3 Acre (House-Driven Value)

Variable	B	Std. Error	Beta	t	Sig.	95% Confidence Interval of B	
						Lower Bound	Upper Bound
Constant	7.562	0.291		25.979	<0.001	6.87	8.136
Residential (Dummy)	-0.199	0.281	-0.28	-0.706	0.481	-0.754	0.357
Commercial and Service (Dummy)	-0.109	0.305	-0.044	-0.357	0.722	-0.712	0.494
Restricted Land (Dummy)	-0.346	0.282	-0.411	-1.227	0.221	-0.904	0.211
Density	0.019	0.014	0.054	1.336	0.183	-0.009	0.046
Sum of Adjacent Land	-0.006	0.008	-0.029	-0.819	0.414	-0.021	0.009
% Commercial / Service	-0.112	0.159	-0.033	-0.704	0.482	-0.425	0.201
%Adjacent Employment	-0.219	0.343	-0.124	-0.637	0.525	-0.896	0.459
%Adjacent Restricted	0.059	0.038	0.065	1.568	0.119	-0.015	0.133
Access to Road (Level)	0.002	0.005	0.016	0.406	0.685	-0.008	0.012
Easy Usable Land	0.023	0.018	0.046	1.24	0.217	-0.014	0.059
Distance to City Core (km)	0.039	0.006	0.241	6.089	<0.001	0.026	0.051
Scaled	-149.028	48.527	-1.107	-3.071	0.002	-244.795	-53.262
Scaled²	6008.531	13862.51	0.33	0.433	0.665	-21348.5	33365.59
Scaled³	261556.8	1156170	0.099	0.226	0.821	-2020096	2543209

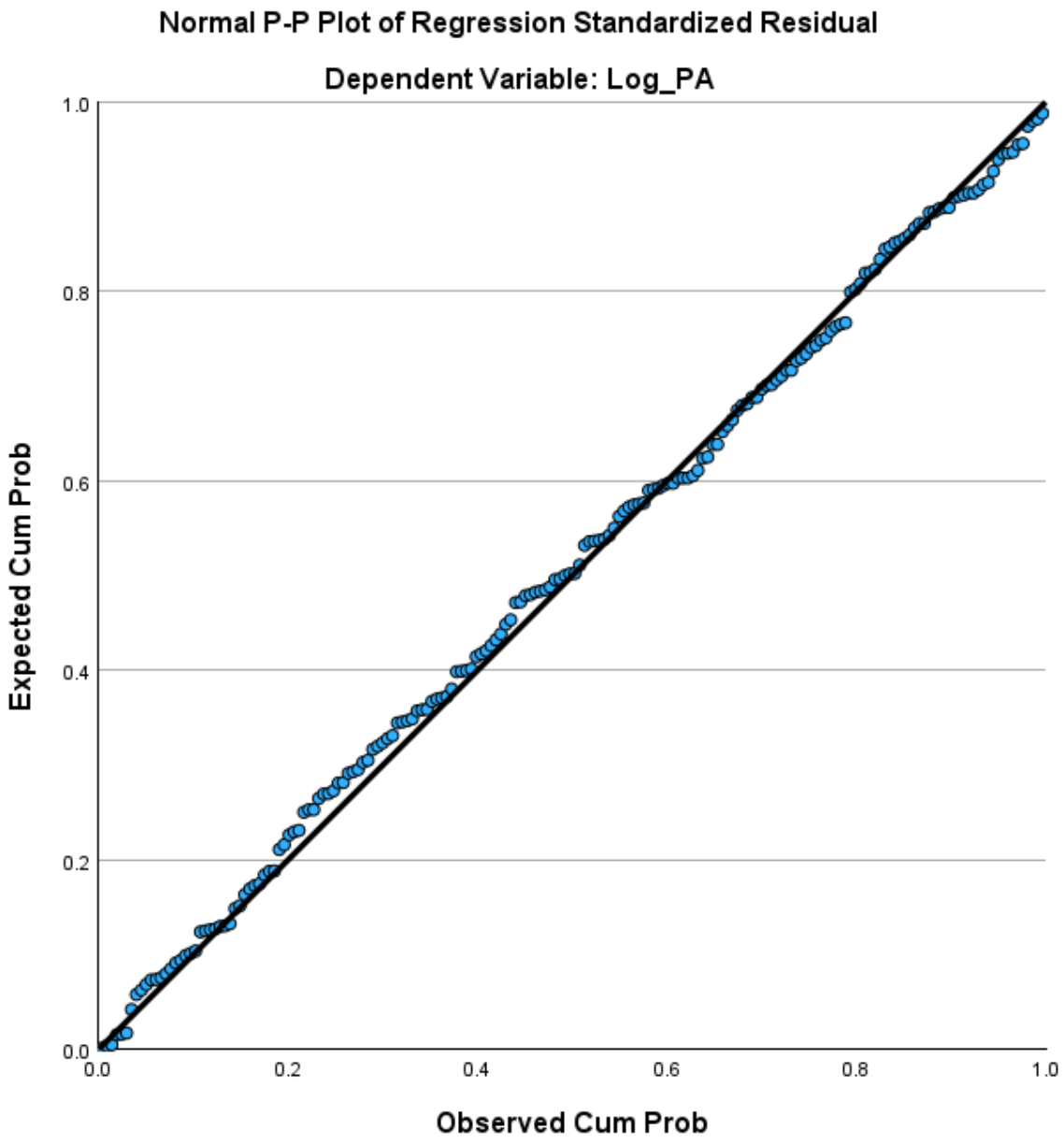


Figure 7 Normal P-P Plot of Regression Standardized Residual from SPSS of Model A: Parcels < 0.3 Acre (House-Driven Value)

Excluded Variable

The variable representing the percentage of adjacent residential land (%Adjacent Residential) was excluded automatically due to perfect multicollinearity, as indicated by a tolerance value of zero. This outcome suggests that the information captured by this variable is fully represented

by other adjacency-related measures included in the model. Its exclusion does not indicate a lack of conceptual relevance but rather reflects redundancy within the adjacent variable set. Furthermore, due to the character of land use designation of house-driven value parcels are more likely to be residential land use, and they often exist as cluster in the city. Which led the percentage of adjacent residential land as excluded variable.

Interpretation

Overall, the results of Model A indicate that land prices for parcels smaller than 0.3 acres are shaped primarily by location-related factors and parcel size effects, while planning-policy variables exhibit limited influence. This pattern reflects the fact that most parcels within this size range are developed residential properties where transaction prices highly reflect the value of existing housing structure rather than underlying land value. The dominance of distance to the city core and the weak performance of ZBL and OP designations support the interpretation that this segment of the market is fundamentally house-driven rather than land-driven. In this context, transaction prices reflect the value of the built structure and neighborhood characteristics rather than the underlying development potential of the land itself.

The magnitude of several coefficients further illustrates how spatial factors influence parcel price within this segment. Because the dependent variable is specified as the logarithm of land price per acre, the coefficients can be interpreted as approximate percentages in value while holding other variables constant. For example, the coefficient associated with distance to the city core (0.039) indicates that land value per acre increases by approximately 4 percent for each additional kilometer farther from the city center. This suggests that small residential

parcels located farther from the central area may command higher prices per acre due to suburban residential preferences and housing market dynamics. Parcel size effects also appear through the scaled parcel-size variable, which is statistically significant and negative. This indicates that even within the small-parcel range, increases in parcel size are associated with lower land value per acre, reflecting the commonly observed inverse relationship between parcel size and unit land price.

These findings align with established literature on small-lot residential markets, which suggests that regulatory signals are not fully capitalized into prices when redevelopment potential is constrained by existing structures. As a result, changes in Official Plan designations or zoning permissions may not be immediately reflected in market prices for small parcels. This reinforces the importance of segmenting parcels by size when evaluating planning-policy impacts and highlights the limitations of applying uniform valuation models across heterogeneous land markets.

4.1.2 Model B: Parcels ≥ 0.3 Acre (Land-Driven Market)

Model Performance

For parcels equal to or larger than 0.3 acre, the regression model demonstrates very strong explanatory power. The model achieved an R value of 0.958, indicating a high degree of correspondence between observed and predicted normalized land prices per acre. The adjusted R^2 of 0.917 confirms that more than 90 percent of the variation in the dependent variable is explained by the included predictors, even after accounting for model complexity. This level of

fitness is substantially higher than that observed in the small-parcel model and reflects the greater sensitivity of development-scale parcels to regulatory, spatial, and contextual factors.

Residual diagnostics further support the reliability of the model specification. Examination of residual distributions indicates approximate normality with limited skewness, and no systematic patterns suggesting heteroscedasticity or model misspecification. These diagnostic results indicate that the regression framework is well suited for capturing land-value dynamics in the large-parcel segment and that coefficient estimates can be interpreted with confidence.

Coefficient Summary

Figure 9 presents the unstandardized coefficients, standard errors, and confidence intervals for Model B. In contrast to the house-driven value model, several planning-policy, adjacency, and physical variables emerge as statistically significant predictors of land value.

Among the planning-policy variables, the restricted land dummy variable exhibits a positive and statistically significant coefficient at the 0.05 level. This result suggests that parcels subject to environmental protection or MTSA-related constraints command higher normalized prices per acre, reflecting the capitalization of scarcity, regulatory certainty, or locational advantages associated with constrained developable land. The residential designation dummy variable, while positive, is not statistically significant, indicating that land-use designation alone does not fully explain price variation without considering additional spatial and regulatory context.

Parcel size effects are also prominent in Model B. The categorical size variable is negative and statistically significant, indicating that, within the land-driven value segment, parcels classified as larger tend to exhibit lower price per acre values, consistent with economies of scale in land

pricing. The scaled continuous parcel size variable is negative and statistically significant, while higher-order polynomial terms (Scale 2, and Scale 3) are weaker and not consistently significant. This pattern suggests diminishing marginal land value per acre as parcel size increases, a relationship that is characteristic of land-driven markets.

Adjacency characteristics play a meaningful role in this model. The sum of adjacent land parcels is negative and statistically significant, indicating that parcels surrounded by a greater number of neighboring parcels tend to exhibit lower land values per acre. This may reflect fragmentation effects or reduced development flexibility in more subdivided contexts. The percentage of adjacent residential land is positive and statistically significant, suggesting that proximity to established residential areas enhances land value, potentially due to perceived market demand or compatibility. In contrast, the percentage of adjacent commercial or service land exhibits a negative and statistically significant coefficient, indicating that adjacent commercial uses may reduce land value for development-scale parcels, possibly due to land-use conflicts or reduced residential desirability.

Distance to the city core emerges as one of the strongest predictors in the model, with a positive and highly statistically significant coefficient. This result indicates that, for larger parcels, normalized land prices per acre increase with greater distance from Aurora's core. This spatial pattern reflects the availability of larger, developable parcels in peripheral areas and aligns with Aurora's growth structure, where greenfield and edge-of-settlement lands often carry greater development potential than centrally located parcels constrained by existing urban form.

Accessibility variables and the proportion of easily usable land do not exhibit statistically significant effects in this model. This suggests that, once zoning permissions, parcel size, and adjacency context are accounted for, variations in road hierarchy and developable land share exert a more limited marginal influence on land value.

Table 7 Coefficients from SPSS of Model B: Parcels ≥ 0.3 Acre (Land-Driven Market)

Variable	B	Std. Error	Beta	t	Sig.	95% Confidence Interval of B	
						Lower Bound	Upper Bound
Constant	6.716	0.206		32.565	<0.001	6.300	7.133
Residential (Dummy)	0.255	0.158	0.236	1.615	0.114	-0.064	0.574
Restricted Land (Dummy)	0.335	0.140	0.317	2.390	0.022	0.052	0.617
Size Category	-0.370	0.133	-0.340	-2.784	0.008	-0.638	-0.102
Density	-0.115	0.066	-0.131	-1.760	0.086	-0.248	0.017
Sum of Adjacent Land	-0.028	0.008	-0.201	-3.518	0.001	-0.043	-0.012
% Adjacent Residential	0.310	0.116	0.200	2.682	0.011	0.077	0.544
% Commercial / Service	-0.611	0.214	-0.177	-2.860	0.007	-1.042	-0.180
Access to Road (Level)	-0.014	0.013	-0.058	-1.034	0.307	-0.040	0.013
Easy Usable Land	-0.119	0.105	-0.069	-1.135	0.263	-0.330	0.092
Distance to City Core (km)	0.135	0.025	0.402	5.486	<0.001	0.085	0.184
Scaled	-6.550	2.473	-2.625	-2.649	0.011	-11.545	-1.556
Scaled²	10.509	5.367	4.057	1.958	0.057	-0.331	21.349
Scaled³	-4.706	3.166	-1.741	-1.486	0.145	-11.100	1.689

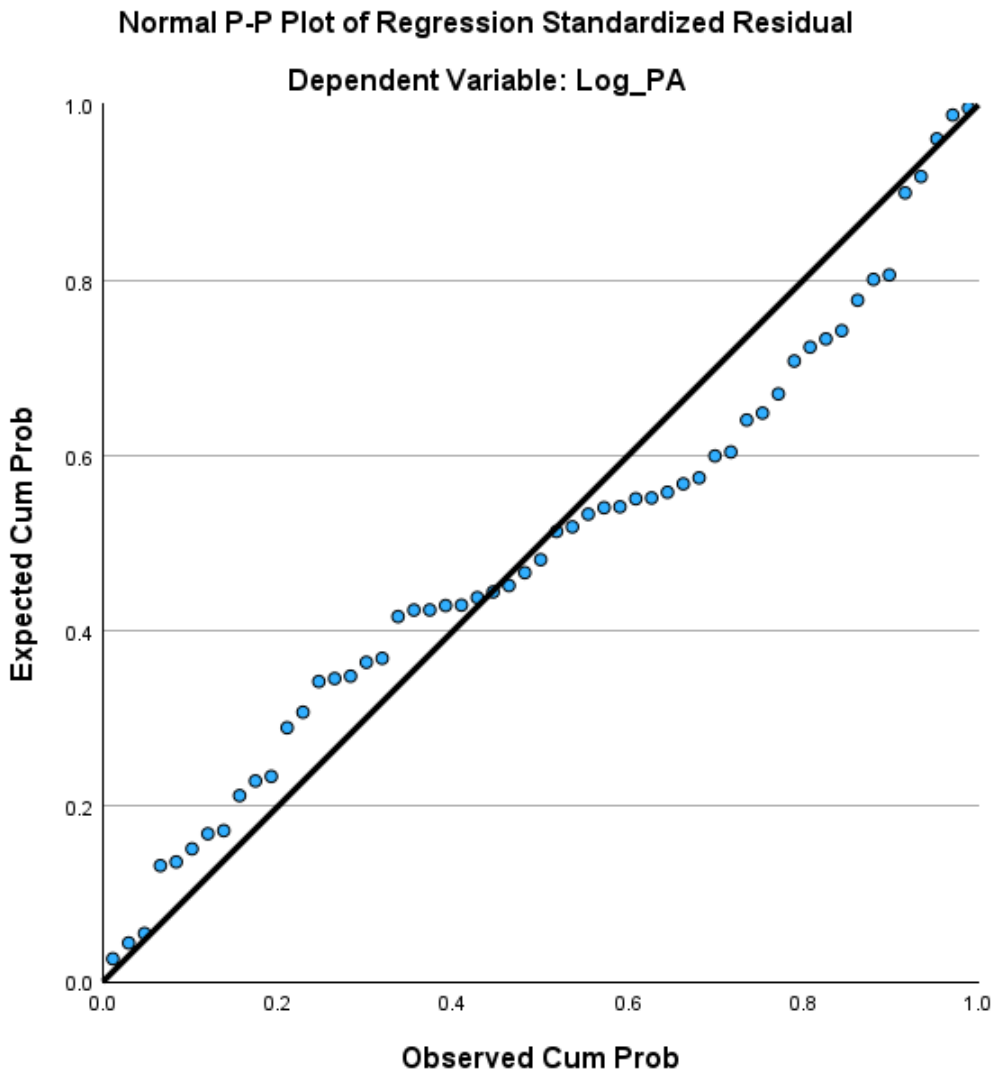


Figure 8 Normal P-P Plot of Regression Standardized Residual from SPSS of Model B: Parcels ≥ 0.3 Acre (Land-Driven Market)

Excluded Variable

The variable representing the percentage of adjacent restricted land (%Adjacent Restricted) was excluded from the model due to perfect multicollinearity, as indicated by a tolerance value of zero. This outcome reflects overlap with the restricted land dummy variable, which directly

captures the presence of environmental or policy constraints on the subject parcel. The exclusion of this variable does not reduce interpretive capacity, as its informational content is already represented in the model.

Interpretation

The results of Model B demonstrate a valuation structure that is fundamentally different from that observed in Model A. In this segment, land values are strongly shaped by planning policy variables, adjacency composition, parcel size, and spatial location. The high explanatory power of the model and the statistical significance of regulatory and contextual variables confirm that parcels equal to or larger than 0.3 acre operate within a land-driven market, where development potential and policy constraints are capitalized into land prices. To detailed classify the impact of parcel area in lands, the categorical parcel size variable was applied to Model B.

The magnitude of several coefficients further illustrates the influence of planning-policy and spatial variables on development-scale parcels. Because the dependent variable is specified as the logarithm of land price per acre, the coefficients can be interpreted as approximate percentage changes in value while holding other variables constant. For example, the coefficient associated with distance to the city core (0.135) indicates that land value per acre increases by approximately 14 percent for each additional kilometer farther from the city center. Similarly, the coefficient for environmentally restricted land (0.335) suggests that parcels containing restricted land exhibit approximately 40 percent higher land value per acre compared with otherwise similar parcels without such constraints. Spatial context also contributes to valuation outcomes. The positive coefficient for the percentage of adjacent residential land (0.310)

indicates that parcels surrounded by higher proportions of residential uses tend to exhibit higher land values, while the negative coefficient for the percentage of adjacent commercial or service land (−0.611) suggests that proximity to commercial uses may reduce land values in development-scale parcels.

These findings indicate that OP designations, zoning allowances, and environmental restrictions play a decisive role in shaping pricing outcomes for development-scale parcels. Unlike small parcels, where the presence of existing structures limits responsiveness to policy change, large parcels reflect market expectations regarding future development opportunities. This distinction underscores the importance of parcel-size segmentation when evaluating planning-policy impacts and provides empirical support for the argument that planning interventions exert their strongest influence in land-driven value contexts.

4.2 Machine Leaning Model Light GBM

In addition to the segmented regression analysis, a LightGBM model was developed to validate the robustness of the regression findings and to assess the predictive importance of planning, spatial, and physical variables. As discussed in Chapter 3, LightGBM is a gradient-boosting algorithm designed to handle nonlinear interactions, multicollinearity, and high-dimensional feature sets more effectively than traditional regression models (Liang et al., 2025). This complementary method provides an opportunity to compare interpretive statistical results with non-parametric predictive outcomes.

The LightGBM model was trained using the same dependent variable as the regression models, namely the log-transformed normalized price per acre, and incorporated all explanatory variables described in Section 3.2.1. The dataset was divided into a training set (80 percent) and a testing set (20 percent), and model hyperparameters were optimized using cross-validation procedures to reduce overfitting risk.

Model Performance

The LightGBM model demonstrated strong predictive performance, outperforming or closely matching the segmented regression models. On the testing dataset, the model achieved an R^2 of 0.778, indicating that 78% of variance in normalized land value was captured by the predictive algorithm. Error metrics including Mean Absolute Error (MAE = 241,425.2) and Root Mean Square Error (RMSE = 0.183) further support the model's predictive reliability (See Table 6).

Residual plots show a tightly centered distribution with no extreme outliers (See Figure 10), and the predicted-vs-true visualization (See Figure 11) indicates strong alignment between observed and modeled prices, with only modest dispersion around the prediction line. However, this is not segmented by 0.3 Acre based on parcel size. Thus, these results confirm the predictive stability of LightGBM and demonstrate its suitability for capturing nonlinear relationships embedded in Aurora's parcel-level land data.

Table 8 LightGBM Model Performance Metrics for Parcel Value Prediction

Metric	Value
Mean Absolute Error (MAE)	241,425.2
Root Mean Squared Error (RMSE)	388,904.6
Root Mean Squared Log Error (RMSLE)	0.184
Normalized Mean Absolute Error	0.0233
Normalized Root Mean Squared Error	0.0376
Normalized Root Mean Squared Log Error	0.0461
Mean Absolute Percentage Error (MAPE)	15.5%
R^2 Score	0.778

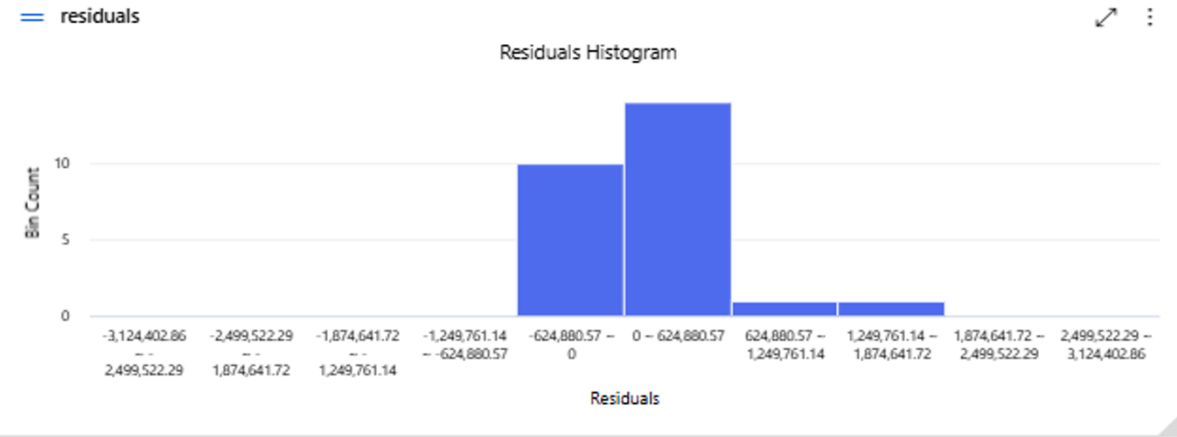


Figure 9 Distribution of Prediction Residuals from the LightGBM Model

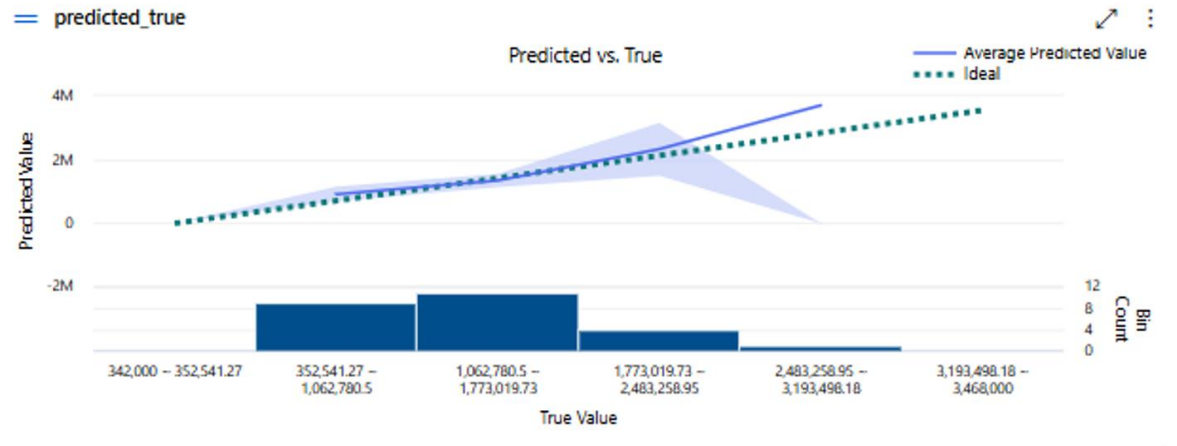


Figure 10 Comparison of Predicted and Observed Parcel Values Using the LightGBM Model (Source: LightGBM model results generated by Elahe Dorani (LandLogic Inc.) using the Aurora parcel dataset.)

Interpretation

The feature-importance profile produced by LightGBM aligns closely with the results of the segmented regression models. Variables related to parcel size, scaled nonlinear terms, adjacency composition, and accessibility emerge as among the most influential predictors. Planning-policy designations—particularly zoning restrictions and OP-related characteristics—also contribute meaningfully to predictive accuracy, especially for parcels with development-scale characteristics.

Because LightGBM does not rely on linearity assumptions, its identification of similar high-impact variables reinforces the robustness of the segmented regression findings and demonstrates that the underlying relationships detected in the statistical models are not artifacts of model specification. Rather, the consistency across both modeling paradigms indicates genuine structural patterns in Aurora's land-value formation processes.

4.3 Comparative Discussion of the Two Segmented Regression Models and the Machine Leaning Model made with LightGBM

This section synthesizes the results of the segmented regression models and the LightGBM analysis to provide an integrated understanding of land-value determinants in Aurora. Together, these methods reveal how parcel size, regulatory designations, adjacency patterns, and accessibility interact to shape land-value dynamics.

4.3.1 Comparison of Model Performance

The segmented regression models showed distinct levels of explanatory power:

- Model A (< 0.3 acre): $R^2 = 0.799$
- Model B (≥ 0.3 acre): $R^2 = 0.917$
- LightGBM Model: $R^2 = 0.778$

The LightGBM model performed comparably or better in predictive accuracy relative to the regression models, which is consistent with its ability to model nonlinear relationships and interactions. The substantially higher R^2 in the ≥0.3-acre regression model confirms that planning-policy and spatial variables exert stronger explanatory influence on land-driven markets.

4.3.2 Comparison of Variable Influence

A comparison of the variable influence across the two segmented regression models and the LightGBM model reveals a consistent divergence in the determinants of land value between small and large parcels in Aurora. In Model A, which captures parcels smaller than 0.3 acre and represents the predominant land composition within the municipality, planning-policy variables exhibit limited statistical significance. This outcome is consistent with the understanding that small parcels are typically valued based on house-driven characteristics rather than underlying land potential. Built-form attributes, which are not explicitly captured in the land-focused modeling framework, dominate buyers' valuation decisions in these contexts. As a result, the influence of zoning, OP designation, and related planning constraints is largely overshadowed,

and the explanatory power of Model A ($R^2 = 0.799$) reflects a more localized and physical-accessibility–driven price structure.

In contrast, Model B reveals a markedly different pattern of influence. Among parcels equal to or larger than 0.3 acre, planning-policy variables such as zoning, environmental restrictions, and OP designations exert substantial influence on land value. The high explanatory power of this model ($R^2 = 0.917$) reflects the extent to which development-scale parcels respond to regulatory permissions and constraints. Larger parcels inherently hold potential for redevelopment or intensification, making planning-policy instruments central components of their market valuation. Adjacency characteristics also exert stronger influence in this segment, suggesting that larger parcels are more sensitive to the compatibility and land-use context of surrounding areas.

The LightGBM model provides an additional layer of insight into variable influence. Although the algorithm does not rely on parametric assumptions, its feature-importance ranking highlights the same groups of variables that appear influential in Model B, particularly those related to parcel size, adjacency, accessibility, and planning constraints. The predictive performance of LightGBM ($R^2 = 0.778$) is closer to Model A than Model B. This outcome is attributable to the distribution of the sample, in which 206 out of 278 observations—approximately 89 percent—fall below the 0.3-acre threshold. The machine-learning model is therefore influenced predominantly by characteristics of the small-parcel market. Nevertheless, the prominence of planning and spatial variables in the feature-importance ranking indicates that, despite sample imbalance, LightGBM captures important structural relationships that become more apparent within the land-driven parcel group.

Together, these findings demonstrate a clear differentiation in the role and influence of key determinants across parcel types, with planning-policy variables emerging as major drivers of land value only in parcels with development-scale potential.

4.3.3 Integrated Interpretation

When interpreted collectively, the segmented regression and LightGBM models provide a coherent and internally consistent understanding of land-value formation in Aurora. The evidence from both approaches indicates that parcel size fundamentally shapes the mechanisms through which land value is determined. In the small-parcel market, which constitutes most of the Aurora's urban fabric, land prices are dominated by the characteristics of existing structures and micro-scale locational attributes. Planning-designation variables register minimal influence in this segment, as the market does not typically treat these parcels as development assets but rather as final-use residential units.

In contrast, the market behavior for parcels equal to or larger than 0.3 acre reflects a distinctly different value structure, one in which planning-policy instruments, environmental restrictions, and adjacency conditions carry significant explanatory weight. The very high R^2 of Model B (0.917) underscores the degree to which land-driven markets respond to regulatory frameworks, environmental constraints, and development potential. The greater sensitivity of large parcels to policy and spatial conditions aligns with established land-economics literature, which suggests that redevelopment-scale parcels internalize regulatory signals more directly into pricing behavior.

The LightGBM results reinforce this interpretation. Although the overall predictive strength of the model is slightly lower ($R^2 = 0.778$), the consistency of its top-ranked variables with those found significant in Model B confirms that the underlying relationships in the data are not artifacts of model specification. Rather, LightGBM replicates key patterns found in the segmented regressions, despite the disproportionate representation of small parcels in the dataset. The method's sensitivity to nonlinear interactions allows it to detect the complex ways in which planning constraints, adjacency patterns, and parcel geometry influence land value, even when the majority of observations belong to a house-driven market.

Taken together, the convergence of findings across models indicates that planning-policy variables meaningfully affect land value only when parcels possess development-scale potential, whereas small parcels reflect a valuation structure dominated by built-form characteristics. This integrated interpretation highlights the importance of model segmentation and demonstrates that the interaction between parcel size, policy environment, and spatial context is central to understanding land-value dynamics in Aurora. The methodological approach employed in this study, which combining segmented regression with machine learning, provides a robust analytical framework for evaluating planning-policy impacts across heterogeneous land markets.

Chapter 5 Interpretation and Discussion

5.1 Valuation of Planning Policy Change's Impact on Land Price

The findings of this research demonstrate that planning-policy changes have uneven but measurable impacts on land prices in the Town of Aurora, with the magnitude of influence varying substantially across parcel sizes. The segmented regression results reveal that parcels equal to or larger than 0.3 acre exhibit strong price sensitivity to OP designations, zoning categories, environmental overlays, and adjacency conditions. These parcels embody development-scale potential, and consequently, regulatory permissions and constraints become integral to market valuation. Model B's high explanatory power underscores this dynamic, indicating that planning-policy variables constitute core determinants of land-driven value.

By contrast, the model for parcels smaller than 0.3 acre indicates limited influence of planning-policy variables. Pricing in this segment is dominated by characteristics linked to the existing built form, such as structural quality, building age, and housing demand patterns, which are not explicitly represented in land-based datasets. This divergence aligns with established literature in urban economics, which suggests that regulatory effects are strongest in markets where land value is separable from building value and where redevelopment potential exists (Ecker & Isakson, 2005; Freemark, 2020). The results demonstrate that regulatory signals embedded in the OP and ZBL become fully capitalized only when market actors perceive development feasibility.

This study therefore contributes empirical evidence that planning-policy changes in Aurora exert their effects primarily within development-scale parcels. The models confirm that instruments such as density permissions, land-use re-designations, and environmental constraints generate clear price signals in larger parcels, suggesting that planning policies can meaningfully reshape market expectations and guide spatial development patterns when the underlying parcels hold sufficient capacity for transformation.

5.1.1 Contribution to Digital Planning Platform and City Service in Aurora

The analytical framework and empirical findings developed in this research directly support the ongoing modernization of Aurora's digital planning services and contribute to the advancement of integrated municipal data platforms. The modeling process, which merges parcel-level geometric data, planning-policy layers, adjacency characteristics, and market transactions, demonstrates how diverse datasets can be harmonized to evaluate the spatialized effects of planning-policy instruments.

These results have implications for the implementation of digital planning platforms such as LandLogic and other emerging decision-support tools that integrate zoning, Official Plan data, land-value models, and spatial analytics. The ability to quantify and visualize how planning-policy changes influence land price provides municipal staff with actionable insights for growth management, development review, and long-term capital planning. The segmented modeling （See equation 3.3.3.1) approach also enhances interpretation by distinguishing between

house-driven and land-driven markets, enabling planners to evaluate policy effectiveness in a more targeted and contextually appropriate manner.

Furthermore, the findings can inform the refinement of City Service Mapping, development charge calculations, and municipal infrastructure prioritization. When planning-policy changes are shown to elevate land values in specific policy areas or corridors, this information can support more precise forecasting of future development activity, guide infrastructure investment sequencing, and promote evidence-based decision making. The methodological approach aligns with the growing emphasis on data-driven planning and supports Aurora's strategic objective of enhancing the transparency, efficiency, and analytical capacity of its planning services.

5.1.2 Generalization to Other Municipalities and Secondary Benefits

The modeling framework developed in this thesis is transferable to other Ontario municipalities and has the potential to generate significant analytical and administrative benefits. The integration of parcel-level valuation, planning-policy layers, and spatial context can be replicated in municipalities that maintain GIS datasets for ZBL, OP, and property information. The ability to systematize and quantify regulatory impacts aligns with broader provincial trends toward planning digitalization, including the Planning Data Catalogue initiatives and the Ontario Land Tribunal's emphasis on data-supported submissions.

Municipalities with growth pressures or undergoing Official Plan updates may benefit particularly from adopting a similar modeling structure. By quantifying the relationship between

policy changes and land-value formation, municipal staff can evaluate the implications of proposed land-use re-designations, density allowances, or environmental overlays before such policies are enacted. This supports more informed scenario testing, improved fiscal impact analyses, and more transparent public engagement processes. Moreover, the predictive capability of LightGBM and other machine-learning models enables rapid screening of regulatory alternatives, which can enhance development review processes and support a more efficient planning approvals pipeline.

The methodological contribution extends beyond valuation. Replication of these models across municipalities builds a basis for regional comparison of planning-policy effects, supports harmonization of planning standards, and encourages consistent and evidence-based decision making across jurisdictions. As municipalities seek to adopt Smart City frameworks and digital solutions for planning, models of this kind provide essential analytical infrastructure for long-term urban policy innovation.

5.2 Limitations

Although the analytical framework developed in this study offers valuable insights into the relationship between planning policies and land values, several limitations must be acknowledged. The first relates to data composition. Because the majority of parcels in Aurora are smaller than 0.3 acre, the dataset is heavily skewed toward house-driven valuation contexts. As a result, the LightGBM model's predictive behavior is influenced predominantly by the small-parcel segment, even though planning-policy impacts are most visible in development-scale

parcels. This imbalance limits the ability of machine-learning models to fully express the influence of planning-policy variables. Furthermore, the limited availability of historical digital OP and planning-policy datasets restricted the development of a consistent time-series framework, thereby limiting the ability to capture long-term temporal response to policy change.

Another limitation relates to sample scalability and data-processing constraints. Several independent variables, including adjacency composition and road-access classification, required substantial manual spatial processing and verification. Due to the limitation of QGIS processing capacity and the intensive time required for parcel-level validation, expanding the analytical sample beyond 300 parcel was not feasible within the study timeline. Consequently, the final sample size reflects a balance between analytical depth, data accuracy, and practical research constraints.

A second limitation concerns the absence of structure-level attributes in the dataset. Small-parcel valuations in Aurora are strongly shaped by building features, including dwelling size, architectural quality, age, and interior condition. These attributes are not represented in municipal planning datasets and were therefore outside the scope of the current model. Their omission constrains explanatory power in the small-parcel model and influences the relative magnitude of planning-policy variables.

A third limitation involves the generalizability of polynomial parcel-size effects. While the cubic transformation of parcel size captures nonlinear behavior effectively within Aurora, the same function may not apply uniformly across different market contexts or cities with distinct urban

forms. Finally, regression-based analyses inherently rely on assumptions regarding linearity, normality of residuals, and independence of errors. Although diagnostic checks support the overall robustness of the models, these assumptions impose structural constraints that do not apply to machine-learning models, which operate more flexibly.

5.3 Further Implications

Despite these limitations, the results of this study offer several broader implications for urban planning practice, land-use governance, and future research. The identification of parcel size as a decisive factor in determining how planning policies translate into market value suggests that municipalities should consider size-based segmentation when evaluating policy impacts. Regulatory interventions, including density changes or land-use re-designations, may require differentiated analysis depending on the development capacity and market context of each parcel.

The findings also underscore the importance of structuring planning-policy datasets in ways that facilitate integration with real estate and spatial data. As municipalities increasingly transition toward digital planning systems, establishing standardized and interoperable data formats will be essential for supporting advanced valuation and scenario-testing models. Future iterations of digital planning tools may benefit from incorporating automated hedonic models or machine-learning–based valuation layers, enabling planners to assess the implications of policy changes in real time.

In addition, the methodological approach developed in this research provides opportunities for expanding the analytical scope in future studies. Incorporating structure-level attributes, temporal price trends, or parcel-level development histories could further refine the understanding of land-value formation. Expanding the model to include multiple municipalities would also enable comparative policy evaluation and support the development of region-wide digital planning practices.

Overall, the findings suggest that planning-policy effects on land value are both measurable and directional, but only when parcels possess sufficient scale for development potential to be realized. This insight contributes meaningfully to planning scholarships and supports the continued advancement of digitally enabled evidence-based land-use planning.

Chapter 6 Conclusions and Recommendations

6.1 Summary of Findings

This research examined the relationship between planning policy change and land valuation in the Town of Aurora through an integrated analytical framework combining segmented regression and machine learning approaches. Rather than treating land price responses as uniform across the market, the analysis reveals that parcel size fundamentally conditions how planning policies are translated into market value. Parcel size therefore functions not merely as a control variable, but as a lens through which planning signals become interpretable.

Parcels smaller than 0.3 acre display valuation behavior characteristic of a house driven market. In this context, planning policies do not fail to matter; instead, their effects are largely masked by the dominance of existing structures and localized housing attributes. Price variation in this segment is shaped primarily by accessibility, neighborhood context, and building related factors that are not directly observable in parcel level planning data. As a result, regulatory signals are absorbed into the housing market and do not emerge as distinct land value effects.

In contrast, parcels equal to or larger than 0.3 acre exhibit a qualitatively different valuation logic. For these parcels, planning designations, environmental constraints, adjacency conditions, and spatial attributes become central determinants of value because they directly define feasible development outcomes. The analysis indicates that regulatory frameworks are capitalized into land price only when parcels possess sufficient scale to convert planning

permissions into economic opportunity. This finding underscores the role of development capacity as the mechanism linking planning policy to market valuation.

The LightGBM analysis reinforces this interpretation from a complementary perspective. Even though the underlying dataset is dominated by small parcels, the machine learning model consistently identifies planning and spatial variables as influential when sufficient development capacity is present. This convergence between interpretable regression results and pattern recognition from machine learning suggests that the observed segmentation is not an artifact of model specification, but a structural feature of Aurora's land market.

Overall, the findings indicate that planning policy changes generate meaningful price signals only under specific market conditions. Planning interventions shape land values most clearly where regulatory permissions can be exercised, while their influence remains indirect and obscured in structurally constrained parcels. This insight reframes planning policy not as uniformly effective or ineffective, but as context dependent, highlighting the importance of parcel scale in evaluating planning outcomes in suburban municipalities such as Aurora.

6.2 Summary of Research Contribution

This study contributes new knowledge to the field of land-use planning by providing an empirically grounded framework for evaluating the impact of planning-policy changes on land valuation at the parcel level. The integration of Official Plan designations, zoning categories, adjacency measures, and spatial accessibility within a unified modeling system offers a methodological advance over traditional land-value studies that rely solely on broad market

indicators or coarse geographies. The development of a segmented modeling approach adds conceptual clarity by distinguishing between house-driven and land-driven markets and by demonstrating that planning-policy instruments operate differently across these contexts.

The research also advances the practice of digital planning within municipal environments. By demonstrating how planning-policy layers, GIS data, and real estate transactions can be harmonized to estimate local land-value responses, this study provides an analytical foundation for future digital decision-support tools. The methodological structure is replicable and thus supports evidence-based planning in municipalities across Ontario that seek to modernize their planning processes and advance geospatial intelligence. The incorporation of machine learning into the analysis further enhances the interpretive robustness of the model and indicates the value of hybrid approaches that combine statistical inference with predictive modeling

6.3 Future Research

The findings of this thesis open multiple pathways for future research that can enhance the understanding of planning-policy impacts on land value and expand the capabilities of digital planning systems. One direction involves the incorporation of additional property-level data, particularly building characteristics that influence prices in house-driven markets. Since small parcels constitute a significant proportion of Aurora's urban composition, integrating structure-related variables such as building age, dwelling size, and renovation quality would improve model accuracy and enable more comprehensive assessment of land-price dynamics in mixed markets.

Another direction for future research involves the exploration of more advanced machine-learning architectures capable of capturing spatial dependencies with greater precision. While LightGBM provides high predictive performance and interpretable feature rankings, it treats spatial relationships implicitly rather than explicitly. Emerging methods in spatial data science, particularly graph-based machine-learning models, offer promising opportunities for capturing networked and relational attributes in land-value formation.

6.3.1 Adding Nearby Property Price

The integration of nearby property sale prices is a particularly relevant extension. Hedonic valuation models and empirical real estate studies consistently show that the prices of neighboring properties exert significant spatial spillover effects on individual parcel values (Anselin 1988; Pace et al. 1998). Incorporating such price spillovers would enable more refined modeling of neighborhood effects and may further clarify the relative influence of adjacency characteristics versus local built-form conditions. This enhancement would improve the model's capacity to differentiate between micro-scale and macro-scale market forces and strengthen its application to both residential and development-scale parcels.

6.3.2 Graph Neural Network

Graph Neural Networks represent another promising avenue for methodological advancement. These models are designed to capture relationships embedded in network or graph structures

and are increasingly applied in spatial analysis and urban studies. A graph representation of parcels, where nodes represent individual parcels and edges represent adjacency or functional relationships, would allow the model to learn spatial dependencies directly rather than relying on manually constructed adjacency features. Recent studies indicate that graph-based deep learning can significantly outperform traditional machine-learning models when analyzing urban datasets with strong spatial structure (Zhang et al. 2023) [new reference]. Applying Graph Neural Networks to Aurora's planning and land-value datasets could reveal higher-order spatial and policy interactions that are not detectable in linear or tree-based models.

6.4 Conclusions

This research demonstrates that land-value responses to planning-policy changes are shaped not only by regulatory design but also by the physical, spatial, and functional heterogeneity of urban land markets. Rather than producing uniform valuation effects, planning interventions interact with parcel characteristics such as size, surrounding land use context, and redevelopment potential to generate differentiated market responses. These findings challenge simplified valuation approaches that treat regulatory changes as having consistent impacts across all parcels and underscores the importance of accounting for underlying market structure when evaluating planning-policy outcomes.

A central contribution of this thesis lies in its explicit distinction between land-driven and house-driven valuation regimes. By demonstrating that planning-policy signals are largely obscured in small residential parcels but become clearly capitalized in development-scale parcels, the study

provides empirical evidence that development capacity functions as a key mechanism linking planning regulation to land value. This insight advances existing literature by moving beyond the detection of price capitalization to explain the conditions under which regulatory effects become economically meaningful.

Methodologically, this research contributes to a spatially explicit and policy-sensitive valuation framework that integrates parcel-level planning data, adjacency conditions, and transaction records within an interpretable empirical modeling structure. The use of segmented regression allows regulatory effects to be examined in a manner that reflects heterogeneous market behavior, while the incorporation of machine-learning models serves to validate nonlinear relationships and reinforce the robustness of the analytical findings. Together, these approaches demonstrate how traditional econometric interpretation and data-driven pattern recognition can be combined to enhance land valuation analysis.

From a planning practice perspective, the findings highlight the growing importance of integrating high-resolution data and analytical tools into municipal decision-making. As municipalities across Ontario continue to modernize planning systems and adopt digital platforms, evidence-based valuation frameworks such as the one developed in this thesis can support more transparent evaluation of policy alternatives, improve understanding of regulatory trade-offs, and inform land-use decisions under conditions of growth and constraint. Finally, the methodological foundation established in this study provides a basis for future research that may extend the framework to other jurisdictions, incorporate dynamic policy effects, or further explore the interaction between planning regulation and market behavior in evolving urban contexts.

While this thesis adopts segmented regression and machine-learning validation to balance interpretability and predictive insight, certain methodological limitations remain. In particular, the segmentation of parcels based on a predefined size threshold relies on theoretical and empirical justification but remains externally specified. Future research could explore the use of decision tree–based methods to endogenously identify parcel segmentation rules and interaction structures within the data. Decision trees may offer additional flexibility in uncovering nonlinear thresholds and hierarchical relationships among parcel size, planning regulations, and spatial context, potentially improving the automatic identification of land-driven and house-driven valuation regimes. Incorporating such approaches could further enhance the adaptability and explanatory depth of parcel-level land valuation models.

References

Ai, H., Liu, Q., Jiang, Y., & He, J. (2020). Urban Residential Land Price Appraisal via Quantifying Impact Factors Based on Deep Belief Networks. *Proceedings of the 2020 12th International Conference on Machine Learning and Computing, ICMLC '20*, 29–33. https://doi.org/10.1145/3383972.3384017

Alonso. (1964). Location and Land use: Toward a general theory of land rent. In *Location and Land Use: Toward a General Theory of Land Rent* (pp. 101–142). Harvard University Press. https://www.degruyterbrill.com/document/doi/10.4159/harvard.9780674730854.c8/html

Ayalke, Z., & Sisman, A. (2022). Nominal Land Valuation with Best-Worst Method Using Geographic Information System: A Case of Atakum, Samsun. *ISPRS International Journal of Geo-Information*, *11*(4), Article 4. https://doi.org/10.3390/ijgi11040213

Bencure, J. C., Tripathi, N. K., Miyazaki, H., Ninsawat, S., & Kim, S. M. (2019). Development of an Innovative Land Valuation Model (iLVM) for Mass Appraisal Application in Sub-Urban Areas Using AHP: An Integration of Theoretical and Practical Approaches. *Sustainability*, *11*(13), Article 13. https://doi.org/10.3390/su11133731

Clauretie, T. M., & Li, H. (2019). Land Values: Size Matters. *The Journal of Real Estate Finance and Economics*, *58*(1), 80–110. https://doi.org/10.1007/s11146-017-9628-x

Demetriou, D. (2016). The assessment of land valuation in land consolidation schemes: The need for a new land valuation framework. *Land Use Policy*, *54*, 487–498. https://doi.org/10.1016/j.landusepol.2016.03.008

Droj, G., Kwartnik-Pruc, A., & Droj, L. (2024). A Comprehensive Overview Regarding the Impact of GIS on Property Valuation. *ISPRS International Journal of Geo-Information*, *13*(6), Article 6. https://doi.org/10.3390/ijgi13060175

Ecker, M. D., & Isakson, H. R. (2005). A unified convex–concave model of urban land values. *Regional Science and Urban Economics*, *35*(3), 265–277. https://doi.org/10.1016/j.regsciurbeco.2004.04.003

Freemark, Y. (2020). Upzoning Chicago: Impacts of a Zoning Reform on Property Values and Housing Construction. *Urban Affairs Review*, *56*(3), 758–789. https://doi.org/10.1177/1078087418824672

Fujita, M. (1989). Urban Economic Theory. *Cambridge Books*. https://ideas.repec.org//b/cup/cbooks/9780521346627.html

Gore, T., & Nicholson, D. (1991). *Models of the Land-Development Process: A Critical Review—T Gore, D Nicholson, 1991*. https://journals-sagepub-com.proxy.lib.uwaterloo.ca/doi/abs/10.1068/a230705

Greenaway-McGrevy, R., Pacheco, G., & Sorensen, K. (2021). The effect of upzoning on house prices and redevelopment premiums in Auckland, New Zealand. *Urban Studies*, *58*(5), 959–976. https://doi.org/10.1177/0042098020940602

Han, W., Zhang, X., & Zheng, X. (2020). Land use regulation and urban land value: Evidence from China. *Land Use Policy*, *92*, 104432. https://doi.org/10.1016/j.landusepol.2019.104432

Hersperger, A. M., Thurnheer-Wittenwiler, C., Tobias, S., Folvig, S., & Fertner, C. (2022). Digitalization in land-use planning: Effects of digital plan data on efficiency, transparency

and innovation. *European Planning Studies*, *30*(12), 2537–2553. https://doi.org/10.1080/09654313.2021.2016640

Hodge, G., Gordon, D. L. A., & Shaw, P. (2014). *Planning Canadian Communities* (7th ed.). Top Hat.

Irwin, E. G., & Bockstael, N. E. (2003). Interacting Agents, Spatial Externalities and the Evolution of Residential Land Use Patterns. In *The Economics of Land Use*. Routledge.

Jaeger, W. K. (2006). The effects of land-use regulations on property values. *Environmental Law*, *36*(1), 105–131.

Jafary, P., Shojaei, D., Rajabifard, A., & Ngo, T. (2024). Automated land valuation models: A comparative study of four machine learning and deep learning methods based on a comprehensive range of influential factors. *Cities*, *151*, 105115. https://doi.org/10.1016/j.cities.2024.105115

Jalali, D., MacDonald, H., Fini, A. A. F., & Shi, S. (2022). Effects of planning regulations on housing and land markets: A system dynamics modeling approach. *Cities*, *126*, 103670. https://doi.org/10.1016/j.cities.2022.103670

Joseph, Q. (2025, December 14). *Newmarket's property assessment up more than $173M*. NewmarketToday.Ca. https://www.newmarkettoday.ca/local-news/newmarkets-property-assessment-up-more-than-25m-11618913

Kauko, T., & d'Amato, M. (2009). *Mass Appraisal Methods: An International Perspective for Property Valuers*. John Wiley & Sons.

Kendall, R., & Tulip, P. (2018a). *Online Appendix: The Effect of Zoning on Housing Prices*.

Kendall, R., & Tulip, P. (2018b). *The Effect of Zoning on Housing Prices*.

Kok, N., Monkkonen, P., & Quigley, J. M. (2014). Land use regulations and the value of land and housing: An intra-metropolitan analysis. *Journal of Urban Economics*, *81*, 136–148. https://doi.org/10.1016/j.jue.2014.03.004

Land Values: Size Matters | The Journal of Real Estate Finance and Economics. (n.d.). Retrieved November 5, 2025, from https://link.springer.com/article/10.1007/s11146-017-9628-x

Liang, P., Cui, X., Ma, J., Song, W., & Xu, Y. (2025). Exploring the Determinants of Spatial Vitality in High-Speed Rail Station Areas in China: A Multi-Source Data Analysis Using LightGBM. *Land*, *14*(6), 1262. https://doi.org/10.3390/land14061262

Mandelker, D. R. (1976). The Role of the Local Comprehensive Plan in Land Use Regulation. *Michigan Law Review*, *74*(5), 899–973. https://doi.org/10.2307/1287830

marketer, S. K. F. B. S. is a content, writer, University, fact-checker S. holds a B. of S. in F. degree from B. S., & policies, helps develop content strategies L. about our editorial. (n.d.). *Understanding MLS: Definition, Benefits, and Fees in Real Estate*. Investopedia. Retrieved November 3, 2025, from https://www.investopedia.com/terms/m/multiple-listing-service-mls.asp

MPAC, M. (2025). *Property Assessment and Property Taxes | MPAC*. https://www.mpac.ca/en/UnderstandingYourAssessment/PropertyAssessmentandPropertyTaxes

Ontario Professional Planner Institute. (2018). *Impact of Land Use Planning on Property Assessment*. Ontario Professional Planners Institute. https://ontarioplanners.ca/blog/planning-exchange/may-2018/impact-of-land-use-planning-on-property-assessment

Rathnayaka, R. M. H., & Prasada, D. V. P. (2020). Impact of Zoning and Archaeological Legislation on Land Values in Kandy Municipal Council Area: A Hedonic-type Analysis. *Tropical Agricultural Research*, *31*(3). https://doi.org/10.4038/tar.v31i3.8396

Robinson, J., Harrison, P., Shen, J., & Wu, F. (2021). Financing urban development, three business models: Johannesburg, Shanghai and London. *Progress in Planning, Financing Urban Development, Three Business Models: Johannesburg, Shanghai and London*, *154*, 100513. https://doi.org/10.1016/j.progress.2020.100513

Stewart, K. G. (2022). How important are land values in house price growth? Evidence from Canadian cities. *Canadian Journal of Economics/Revue Canadienne d'économique*, *55*(1), 249–271. https://doi.org/10.1111/caje.12585

Town of Aurora. (2025). *Official Plan*. Town of Aurora. https://www.aurora.ca/business-and-development/official-plan/

Town of Aurora. (2026a). *Community Profile and Site Selection Information*. Town of Aurora. https://www.aurora.ca/business-and-development/community-profile-and-site-selection-information/

Town of Aurora. (2026b). *Town of Aurora Offical Plan*.

Wu, L., Zhang, Y., Wei, Y., & Chen, F. (2022). A BP Neural Network-Based GIS-Data-Driven Automated Valuation Framework for Benchmark Land Price. *Complexity*, *2022*(1), 1695265. https://doi.org/10.1155/2022/1695265

Wyatt, P. (2022). *Property Valuation*. John Wiley & Sons.