

# Predicting NIMBYism

Daniel Lewis

dl3645@columbia.edu

Columbia University

Urban Planning

December 10, 2025

## 1 Introduction

This thesis investigates the patterns of neighborhood opposition to housing development in Austin, Texas, and evaluates the democratic implications of using predictive analytics to anticipate such behavior. The project asks: how can cities identify and understand patterns of neighborhood opposition, and what are the normative consequences of using algorithmic tools to manage this participation?

Austin, Texas faces an unprecedented housing affordability crisis that has fundamentally reshaped its political landscape. Between 2010 and 2020, Austin’s population grew by 33%, adding nearly 250,000 residents (**census2020**). This growth, driven primarily by the technology sector, created severe supply constraints. By May 2022, the median home price reached \$550,000, up 79% from January 2020. As of August 2025, prices have moderated to \$444,490 but remain unaffordable for most residents (**unlockMLS2025**).

The research addresses a critical gap at the intersection of urban planning and algorithmic governance. As cities turn to technological solutions to streamline development, the use of

predictive tools raises questions about democratic participation and civil liberties. The findings will directly inform policy in Austin—where Development Services could use insights to improve engagement—and establish ethical frameworks for other cities.

### **Research Questions**

The project is guided by three questions:

- a. What demographic, geographic, and behavioral factors predict whether residents will oppose specific development projects?
- b. How do different stakeholders perceived the legitimacy and utility of algorithmic tools in housing policy?
- c. What alternative technological approaches could improve development processes without raising concerns about democratic participation?

An important recent development shapes this work. In 2025, the Texas Legislature passed House Bill 24, raising the valid petition protest threshold from 20% to 60% and eliminating the automatic supermajority requirement (effective Sept 1, 2025). While these changes occurred after our study period (2018-2025), they create crucial context: we are analyzing behavior under a regime that has now been substantially weakened but continues to operate in other states.

## **2 Background**

Austin’s governance structure complicates housing policy implementation. The city operates under a council-manager system where the mayor has limited executive authority. Passing controversial reforms has historically required council supermajorities to overcome property owner protests authorized by Texas law (**cityaustin2010**).

Historically, neighborhood associations have mounted organized opposition to development using tools like valid petitions and public testimony. This opposition often correlated with

neighborhood demographics, raising questions about equity. In response to the deepening crisis, Austin underwent a political transformation. The November 2022 elections delivered a pro-housing supermajority (**tcclerk2022**). This shift enabled sweeping reforms: eliminating parking requirements citywide in November 2023 (**axios2023parking**), reducing minimum lot sizes, and allowing density on corner lots through the HOME Initiative (**pool2023**).

These reforms have generated measurable impacts, with apartment vacancy rates climbing to 10-15% by 2025 (**kvue2024vacancy**). However, implementation challenges persist, with permit volumes fluctuating and staffing pressures in the Development Services Department. Throughout this period, opposition tactics have evolved. Austin’s unique position as a technology hub, home to major tech firms and a workforce of 180,000, provides a distinct civic context for debating the role of data in managing this conflict. This research contributes to literature on algorithmic governance and housing policy by separating the empirical structure of opposition from the normative question of prediction, providing evidence on the tradeoffs between efficiency and legitimacy that informs theoretical debates about technocratic city management.

### **3 Literature Review**

This section organizes the literature around three themes that correspond to the three research questions introduced above. Theme A addresses the determinants of neighborhood opposition to housing development and informs our first research question about which factors should predict opposition. Themes B and C address our second and third research questions by examining the possibilities and limits of applying machine learning to predict political behavior in democratic urban governance.

#### **3.1 Theme A: Homeownership, Local Interests, and Opposition to Housing**

William Fischel’s “Homevoter Hypothesis” (**fischel2001**) provides the foundational theoretical framework for understanding neighborhood opposition to housing development. Fischel argues

that homeowners, whose primary asset is typically their residence, rationally oppose development that might reduce property values or alter neighborhood character. This creates systematic political opposition to housing supply increases, even when broader social welfare would benefit. The hypothesis implies that homeownership status, property values, and length of residence should be positively associated with opposition intensity—predictions that our machine learning model will explicitly test using Austin’s detailed opposition data.

McCabe’s research on homeownership and local political participation extends Fischel’s framework by showing that homeowner-dominated municipalities are more likely to restrict multifamily housing and adopt restrictive land use regulations (**mccabe2016**). In McCabe’s account, the concentration of homeownership within a jurisdiction shapes both policy outputs and the structure of political conflict. Our analysis will test whether variation in homeownership concentration across Austin neighborhoods is associated with the patterns of zoning opposition that McCabe’s theory would predict.

Recent econometric work by Hankinson provides more granular evidence on the determinants of housing opposition using survey experiments across multiple U.S. metropolitan areas (**hankinson2018**). Hankinson finds that even renters in high-cost markets can exhibit NIMBY attitudes when they experience “price anxiety” about future housing costs, particularly when potential new development is sited near their own homes rather than elsewhere in the city. Our machine learning approach can test whether Hankinson’s price-anxiety mechanism operates in Austin by examining whether opposition patterns differ between homeowners and long-term renters in rapidly appreciating neighborhoods. If our predictive model finds that renter opposition increases in high-appreciation areas, it would provide novel support for Hankinson’s theory using revealed-preference behavioral data rather than survey responses.

**How our research tests these theories.** Drawing on Travis Central Appraisal District records, City Council testimony, and tract-level demographic data, we will test whether the key variables highlighted by Fischel (homeownership status, property value, tenure) and McCabe (homeownership concentration) predict opposition behavior in Austin. If these variables emerge as

strong predictors with the expected signs, our results will substantiate these foundational theories in a new empirical context. If other factors dominate—such as proximity to the proposed development, age structure, or past political participation—this would suggest that the existing theories are incomplete or that Austin’s particular institutional and demographic context generates distinct dynamics.

### **3.2 Theme B: Machine Learning, Political Behavior, and Democratic Legitimacy**

The application of machine learning to predict citizen political behavior raises fundamental questions about democratic governance that have only recently received systematic scholarly attention. Kleinberg et al. show that machine learning models can substantially improve prediction of judicial decisions, such as pretrial release and bail, relative to human decision-makers (**kleinberg2018**). At the same time, they emphasize that such tools can change the decisions they are supposed to predict, raising concerns about feedback effects and the normative acceptability of delegating high-stakes judgments to algorithms. This tension is directly relevant to predicting neighborhood opposition: if city officials used an opposition-forecasting model to pre-emptively manage hearings or outreach, the model could shape the very political behavior it is measuring.

Eubanks’ analysis of “automating inequality” in welfare, child protective services, and homeless services provides a critical framework for evaluating algorithmic governance tools (**eubanks2018**). She documents how predictive systems that are framed as neutral risk-management tools often intensify the surveillance and punishment of poor and marginalized populations. Eubanks argues that such systems create “digital poorhouses” that embed existing inequalities in technical infrastructures. Extending this framework to housing policy, we must ask whether a NIMBY prediction algorithm would systematically identify opposition in particular neighborhoods (for example, wealthier and whiter areas) and whether such identification would advantage development in other areas or stigmatize residents who exercise their democratic right to oppose projects.

Barocas and Selbst provide a widely cited legal-technical analysis of how machine learning

can produce discriminatory outcomes even when protected characteristics are not directly used as inputs (**barocas2016**). They show how seemingly neutral variables can function as proxies for race, class, or other protected attributes, leading to “disparate impact” under antidiscrimination law. In the context of NIMBY prediction, variables such as property values, appraisal protest behavior, or neighborhood characteristics could serve as proxies for race or income, generating predictions that disproportionately affect particular groups even without explicit intent.

**How our research advances this literature.** The existing literature demonstrates that machine learning can improve prediction in high-stakes settings while raising serious fairness and legitimacy concerns. However, there is little empirical work that builds and evaluates predictive models of citizen political behavior in land use and zoning contexts while simultaneously examining democratic legitimacy concerns. Our study contributes by (1) constructing a parcel- and owner-level predictive model of zoning opposition, and (2) pairing that technical work with qualitative interviews that elicit stakeholders’ own evaluations of whether such a tool would be legitimate, fair, or democratically acceptable. This dual approach allows us to assess both the technical feasibility and the democratic acceptability of NIMBY prediction, directly informing our second research question.

### **3.3 Theme C: The Limits of Algorithmic Governance in Democratic Contexts**

Beyond fairness concerns, recent scholarship has examined how algorithmic tools reshape the practices and boundaries of democratic governance. Kitchin’s critique of “real-time cities” argues that smart-city dashboards and data-driven management systems tend to prioritize technocratic efficiency and managerial control over democratic accountability (**kitchin2014**). He identifies emerging problems including the politics of urban data, technocratic lock-in, and the risk that optimization logics displace contestation and deliberation. These concerns are salient for any algorithmic tool that would allow city officials to anticipate and manage political opposition.

Schuilenburg and Peeters analyze smart-city security systems as expressions of “pastoral

power,” in which behavioral scripts embedded in urban technologies gently steer citizens toward desired conduct in public space (**schuilenburg2018**). Their case study of Eindhoven’s “De-escalate” project shows how sensor networks and algorithmic interventions can normalize particular uses of space while discouraging others. Although their focus is on security and public order, the broader lesson is that algorithmic systems can subtly reshape who feels welcome to participate in public life and under what conditions.

Zarsky develops an analytic roadmap for evaluating algorithmic decision systems in terms of efficiency, fairness, transparency, and due process (**zarsky2016**). He argues that opaque, automated decisions can undermine individuals’ ability to understand, contest, and appeal outcomes that affect them, especially when decisions are based on complex statistical models. In the housing opposition context, a predictive model that profiles certain owners or neighborhoods as likely opponents could influence how officials treat their testimony or whether they prioritize outreach to particular groups, raising questions about procedural fairness and the right to contest decisions.

**How our research contributes.** Together, this literature suggests that even technically accurate predictive tools can be normatively problematic in democratic governance. Our qualitative interviews with city officials, neighborhood representatives, housing advocates, developers, and civic technologists are designed to probe these concerns directly. We will examine whether stakeholders view NIMBY prediction as a legitimate planning tool, as a form of surveillance that chills democratic participation, or as acceptable only under specific conditions (for example, public transparency about the model and strict limits on how predictions are used). These findings will inform our second and third research questions by identifying the democratic limits of algorithmic governance in housing and by highlighting alternative technological approaches—such as tools for transparency, scenario exploration, or capacity-building—that may improve development processes without profiling or predicting opposition.

## 4 Research Design

This study combines quantitative analysis of existing administrative records with new qualitative interviews and public meeting observations. This design allows us to both identify empirical patterns in opposition behavior and understand how stakeholders evaluate the democratic implications of using predictive tools to anticipate that behavior.

### 4.1 Research Site Selection

Austin, Texas provides an ideal research site for several reasons. First and foremost, Austin is unique among major U.S. metropolitan areas in maintaining comprehensive public records of protest petitions and formal opposition to zoning changes. A public information request survey of the top 50 major metros operating under similar state laws revealed that Austin alone has systematically preserved these records in a form amenable to parcel-level analysis. This exceptional data availability makes possible an empirical analysis that would be infeasible in nearly any other U.S. city.

Second, the study period spans both zoning stasis and dramatic reform. The concentrated reform period in late 2023, anchored by parking requirement elimination on November 2, 2023 and HOME Phase 1 passage on December 7, 2023, creates a natural experiment in how zoning liberalization affects opposition patterns. These reforms were followed by HOME Phase 2 in May 2024, which further expanded allowable density on corner lots and other properties. This variation allows us to examine whether factors predicting opposition remained stable or shifted as the policy regime changed.

Third, Austin’s position as a technology hub with approximately 180,000 tech workers creates both the technical capacity for algorithmic governance and an informed civic discourse about its implications (efa2025; idc2024). This context means stakeholders interviewed will likely have more sophisticated understandings of algorithmic tools than in cities with less civic tech infrastructure.



Finally, Austin’s diverse stakeholder landscape, including active neighborhood associations, organized housing advocacy groups such as Austinites for Urban Rail Action (AURA), and a council-manager government structure requiring supermajorities for controversial reforms, provides rich variation in how different groups might perceive predictive tools.

## 4.2 Quantitative Component

**Data Sources, Linkage, and Management.** The quantitative analysis will construct a parcel-level and owner-level dataset linking multiple administrative and public record sources:

1. **Opposition records (2007–2025):** Protest petitions, valid petitions under Texas Local Government Code § 211.006(d) (pre-HB 24), and formal opposition filings extracted from zoning case files via public information requests.
2. **City of Austin development records (1966–2025, using 2007 onwards):** Site plan cases and zoning review cases from the City of Austin Open Data Portal, including case outcomes, processing timelines, and parcel identifiers.
3. **Travis Central Appraisal District records (2021–2025, optionally earlier years via PIR):** Parcel-level property characteristics including assessed value, land use, and approximate ownership tenure (`tcad`public`).
4. **Demographic context (2000–2025):** U.S. Census Bureau tract-level data including decennial census (2000, 2010, 2020) and American Community Survey annual estimates (2005-forward).
5. **Public testimony (2015–2025):** City Council meeting transcripts and archived ATXN video recordings.
6. **Campaign finance data (2016–2025, optional):** Campaign contribution and expenditure records from the City of Austin Open Data Portal (`austin`campaign`finance`transactions; austin`campaign`finance`reports`).

To maintain confidentiality and follow IRB best practices, we will implement a two-table structure: a **linkage file** (Table A) containing study\_id, names, exact addresses, and other direct identifiers, stored in encrypted, access-controlled files with access limited to the Investigator; and an **analytic dataset** (Table B) containing study\_id and derived variables without direct identifiers, used for all analysis and publication.

**Feature Construction and Variables.** The analytic dataset will include three categories of predictors: *property and ownership characteristics* (assessed value, homeownership status, ownership tenure, property type, appraisal protest history); *spatial variables* (proximity to proposed developments, neighborhood demographics, location relative to council districts and neighborhood association boundaries); and *behavioral and temporal variables* (prior participation in zoning testimony, campaign contributions, policy regime indicators). An additional category will examine **demographic differences between property owners and nearby residents.**

**Modeling Approach and Validation.** We will implement and compare logistic regression as an interpretable baseline and tree-based methods (Random Forests, Gradient Boosted Trees) that can capture nonlinear relationships. Model performance will emphasize metrics appropriate for rare events: precision-recall curves, F1 score for the positive class, calibration diagnostics, and ROC AUC.

To ensure temporal validity and avoid data leakage, we will use an **annual expanding window validation strategy:**

- Year 1: Train on data through 2018, test on 2019, validate on 2020–2025
- Year 2: Train through 2019, test on 2020, validate on 2021–2025
- Year 3: Train through 2020, test on 2021, validate on 2022–2025
- Year 4: Train through 2021, test on 2022, validate on 2023–2025
- Year 5: Train through 2022, test on 2023, validate on 2024–2025
- Year 6: Train through 2023, test on 2024, validate on 2025
- Year 7: Train through 2024, test on 2025, predict 2026 (no validation)
- Year 8: Train through 2025, predict 2026–2027 (no testing or validation)

This mimics real-world deployment with annual retraining on accumulating historical data.

### 4.3 Qualitative Component

**Subject Selection and Sampling.** The qualitative component will involve approximately 10 semi-structured interviews with stakeholders who directly participate in or are affected by Austin’s housing development processes. We will use purposive sampling to recruit a diverse set of participants: 2 city officials/planners from Development Services Department; 2 neighborhood association leaders; 2 housing advocates from organizations such as Austinites for Urban Rail Action (AURA); 2 civic technologists with expertise in civic data or algorithmic governance; and 2 developers (optional). We do not claim to reach thematic saturation; the qualitative component is designed to surface major tensions rather than exhaust the space of views. Snowball referrals may be used to identify additional participants up to this target if needed.

**Interview Protocol.** Interviews will be semi-structured, following the IRB-approved interview guide (see Appendix A), covering: participant’s role and experience; perceptions of current barriers in development approval; awareness of and attitudes toward algorithmic tools; reactions to predictive modeling of opposition; concerns about fairness, transparency, and democratic legitimacy; alternative technological approaches; and conditions for acceptability. Interviews will last 15–30 minutes, conducted in person, by phone, or via secure videoconference.

**Observational Data.** The study will include non-participant observation of public meetings over a six-month period, including live video attendance at City Council and Planning & Zoning Commission meetings and review of archived ATXN recordings. Field notes will document how stakeholders discuss predictive tools, data-driven governance, and democratic participation concerns.

**Qualitative Analysis Plan.** Analysis will follow an iterative thematic coding approach using NVivo or Atlas.ti. Deductive codes will be adapted from Grimmelikhuijsen and Meijer’s (2022) legitimacy framework and procedural justice concepts. Inductive codes will emerge from close reading of transcripts. We will present representative quotations to ground key themes in

participants' own words.

## **4.4 Integration of Quantitative and Qualitative Findings**

The mixed-methods design allows us to triangulate findings. The quantitative component establishes what patterns exist in opposition behavior and which factors predict them. The qualitative component examines how stakeholders interpret the benefits, risks, and legitimacy of using those patterns predictively. Integration will occur at the interpretation stage, synthesizing empirical patterns with normative stakeholder evaluations to address the overarching research question.

## **5 Task Schedule**

### **I. 2025 October–November: Literature Review and Data Acquisition**

- 10/01–11/15: Complete comprehensive literature review
- 11/25: Submit IRB materials (target approval mid-December)
- 10/15–12/15: Acquire datasets via public information requests to Travis Central Appraisal District and City of Austin Open Data Portal
- **Milestone:** Complete thesis proposal (12/10)

### **II. 2025 December – 2026 January: Quantitative Analysis**

- 12/15–01/15: Data preparation, cleaning, and linkage; construct two-table structure
- 01/10–01/25: Feature construction and exploratory analysis
- 01/15–02/10: ML model development with annual expanding window validation
- 01/20–02/10: Geographic and temporal analysis of opposition patterns
- **Deliverable:** Draft quantitative methods and results chapter (02/10)

### **III. 2026 January–February: Qualitative Data Collection**

- 01/05–01/25: Finalize interview protocol, recruit participants via purposive sampling

- 01/20–02/28: Conduct 10 interviews (15–30 minutes each)
- 01/15–03/15: Non-participant observation (live attendance and archived ATXN review)
- 02/01–03/10: Transcribe and begin preliminary coding
- **Deliverable:** Complete interview transcripts and field notes (03/10)
- **Note:** Schedule in-person Austin time during this period, secure Columbia travel funding

#### **IV. 2026 February–March: Integration and Analysis**

- 02/20–03/20: Systematic qualitative coding using Grimmelikhuijsen & Meijer (2022) framework
- 03/01–03/15: Analyze comparative frameworks from other cities
- 03/10–03/25: Integrate quantitative and qualitative findings
- 03/15–03/30: Develop theoretical framework connecting findings to HB 24 context
- **Deliverable:** Complete draft chapters submitted to advisor (03/30)

#### **V. 2026 March–May: Writing and Defense**

- 04/01–04/10: Revise based on advisor feedback
- 04/03: Distribute penultimate draft to reader(s) (14 days before jury, required)
- 04/05–04/15: Prepare jury presentation
- 04/13–04/17: Thesis jury (required for graduation)
- 04/18–05/05: Final revisions incorporating jury feedback
- **Final submission:** 05/08/2026

## **A Interview Guide**

### **Predicting NIMBYism: Stakeholder Interview Guide**

#### **Introduction**

- Thank you for agreeing to speak with me. I am a graduate student at Columbia University studying housing development processes in Austin.

- This interview will last approximately 15–30 minutes.
- Participation is voluntary. You may decline to answer any question or stop at any time.

*If recording:* Do I have your permission to record this conversation for accuracy?

### **I. Professional Background and Context**

1. Could you briefly describe your role and how you interact with housing development or zoning processes in Austin?
2. Based on your experience, what are the most significant factors that currently delay or prevent housing development in Austin?
3. How would you characterize the nature of community engagement or neighborhood opposition to revisions/re-zonings in recent years?

### **II. Perceptions of Predictive Tools**

1. Are you aware of any data-driven or algorithmic tools currently used by the City of Austin (e.g., in permitting, service delivery, or planning) that you find particularly useful or problematic?
2. Cities are increasingly using data to model urban trends. What is your initial reaction to the concept of using data to predict where neighborhood opposition to zoning changes is most likely to occur?
3. What potential benefits, if any, could you see from using such a tool? (For example: better resource allocation, earlier outreach, identifying engagement gaps).
4. What risks or concerns would you have about such a tool? (For example: bias, transparency, fairness, exclusion).

### **III. Democratic Legitimacy and Governance**

1. Do you believe that using predictive analytics to anticipate public participation would make the planning process more or less democratic? Why?
2. If the City were to adopt more predictive technologies for planning, what specific safeguards or transparency measures would be essential for you to trust them?
3. Is there anything else about the intersection of technology, data, and housing policy in Austin that you would like to share?

## **B Recruitment Materials**

**Email Subject:** Interview Request: Research on Austin Housing Policy and Technology

Dear [Name],

I am writing to invite you to participate in a research study regarding housing development processes in Austin. I am a graduate student in Urban Planning at Columbia University, conducting this research for a thesis under the supervision of Dr. Hiba Bou Akar.

**Purpose of the Study:** This study examines patterns of neighborhood opposition to housing development and evaluates how different stakeholders perceive the use of new data-driven and predictive tools in planning. We are seeking perspectives from [City Officials / Neighborhood Leaders / Housing Advocates / Developers / Civic Technologists] to understand diverse views on transparency, efficiency, and fairness in the development process.

**Procedures:** Participation involves a single semi-structured interview lasting approximately 15–30 minutes. The interview can be conducted in person, by phone, or via secure video-conference (e.g., Zoom).

**Voluntary Participation:** Your participation is completely voluntary. You may decline to answer any questions or withdraw from the study at any time without penalty.

If you are willing to share your perspective, please let me know your availability for a brief conversation in the coming weeks.

Thank you for your time and consideration.

Sincerely,

Daniel Lewis Investigator Graduate School of Architecture, Planning and Preservation  
Columbia University dl3645@columbia.edu

## C Public Meeting Observation Protocol

**Date:** \_\_\_\_\_

**Location:** \_\_\_\_\_

**Meeting Type:** ☐ City Council    ☐ Planning & Zoning Commission    ☐ Neighborhood Association

**Format:** ☐ Live attendance    ☐ Archived ATXN video review

### **Attendees (general categories, no names):**

- Approximate number present: \_\_\_\_\_
- Stakeholder groups represented: \_\_\_\_\_

### **Topics Discussed Related to Study:**

- Housing development or zoning matters: \_\_\_\_\_
- References to data, predictive tools, or algorithmic approaches: \_\_\_\_\_
- Discussion of neighborhood opposition or participation: \_\_\_\_\_

**Field Notes:** [Space for detailed notes on how stakeholders discuss housing development processes, predictive tools, democratic participation, and related themes. Focus on publicly observable statements and behavior that are already part of the public record.]

### **Follow-up Questions or Observations:**