Unpacking Housing Supply Constraints: Zoning Regulations and Hidden Barriers

Jaehee Song*

June 26, 2026

Abstract

Housing supply is shaped by land availability, construction costs, regulations, and administrative barriers. Using parcel, zoning, and construction cost data for Greater Boston, I examine how these constraints jointly shape single-family development. I develop a discrete-choice model where landowners choose whether and what to build and estimate that implicit development costs beyond formal zoning add \$117,000 per project at the median, plus \$45,000 for nonconforming development. Counterfactuals show that reducing the per-project development cost by \$50,000 raises supply by 73–79% and removing minimum lot size requirements raises supply by about 27%. Individually these reforms lower prices by 1% or less, but combined they raise supply by 130–140% and lower prices by about 2–4%, indicating that the largest affordability gains come from relaxing multiple constraints together rather than any one in isolation.

JEL Classification: R31, R52, R14

Keywords: housing supply, zoning, regulatory costs, discrete choice

*University of Colorado, Leeds School of Business (email: jaehee.song@colorado.edu).

1 Introduction

Housing prices have increased sharply in many U.S. metros, alongside evidence that a range of supply-side constraints limit the extent to which housing supply adjusts to rising demand (see Baum-Snow and Duranton, 2025, for a review). These constraints include limited developable land due to geography and existing development patterns (Saiz, 2010), rising construction costs and labor shortages (Glaeser and Gyourko, 2025; Howard et al., 2024), and lengthy construction timelines (Mayer and Somerville, 2000; Gabriel and Kung, 2025). In many places, zoning codes restrict the quantity, density, and type of housing that can be built (Gyourko and Molloy, 2015; Jackson, 2016; Gyourko and Krimmel, 2021), while local politics shape both the design of these rules and informal administrative barriers, such as discretionary review and political opposition (Glaeser et al., 2005; Hilber and Robert-Nicoud, 2013; Fang et al., 2023; Favilukis and Song, 2023; Dobbels and Tavakalov, 2024; Mast, 2024). However, existing work has largely studied them in isolation, leaving open the question of how their stringency covaries and how they interact to shape housing supply.

These constraints are intertwined, so the extent to which any one affects housing supply depends on the others. For example, a large minimum lot size may matter little where land is abundant and inexpensive, but it can sharply restrict supply in places where only smaller-lot development would otherwise be profitable. Likewise, lowering regulatory barriers may have limited effects in locations with little developable land or high construction costs. As such, it is difficult to gauge the role of any single constraint, or the effects of relaxing it, without a framework that accounts for how multiple constraints operate together.

In this paper, I quantify how land availability, construction costs, zoning regulations, and implicit development costs shape single-family housing supply in the Greater Boston metropolitan area. I assemble parcel-level data on current and past housing inventories and vacant land, district-level zoning regulations, and construction cost estimates that vary across municipalities and housing types. I then develop and estimate a discrete-choice model of single-family housing construction in which landowners choose whether to develop and what type of home to build, subject to these constraints. I use the model to recover municipality-level implicit development costs that reflect barriers in the permitting and approval process, and the uncertainty around them, beyond those captured by formal zoning rules and physical construction costs. Using the estimated model, I conduct counterfactual analyses to quantify how relaxing supply constraints, separately and jointly, changes the level and composition of housing supply and, in turn, home prices.

I compile several datasets to characterize housing supply constraints at a granular spatial scale. The Metropolitan Area Planning Council (MAPC) Zoning Atlas provides zoning regulations in Greater Boston, including minimum lot sizes, floor-area-ratio restrictions, and other dimensional requirements at the zoning-district level. ATTOM Assessor History and Recorder data provide parcel characteristics and transaction prices, which I use to identify developable land and existing structures and to impute land and structure values through hedonic price models. Craftsman Costbooks provide construction cost estimates by building characteristics and location, which I use to construct municipality-specific costs for the building types commonly observed in each municipality. The final analysis sample comprises 501,959 single-family and vacant parcels across 98 municipalities.

I begin by documenting housing supply conditions and development patterns in the Greater Boston area. Single-family development is rare overall (1.8% of parcels over 2010–2019) but far more common on vacant than on improved parcels (10.5% versus 1.2%). Zoning stringency varies widely, with minimum lot size requirements ranging from 7,500 square feet at the 10th percentile to 45,000 at the 90th, and new construction bunches visibly at the minimum lot size threshold, with about 13% of newly built homes on lots within 5% of the requirement, suggesting that formal zoning laws, specifically minimum lot sizes, bind in this setting. At the same time, smaller-lot building types show much higher profit per square foot of land than larger-lot types, with a median unit profit of $26 per square foot for small-lot types versus $4 for large-lot types, implying that large minimum lot sizes can deter some of the most profitable forms of development. Yet across municipalities, development rates are not highest where observed profit margins are largest, with the raw correlation between the two in fact negative (-0.19). This is consistent with high-price municipalities tending to have less vacant land and facing stricter implicit barriers. Therefore, it underscores the difficulty of interpreting any single supply-side factor alone and motivates a framework that jointly accounts for land availability, construction costs, zoning, and implicit barriers.

Next, I develop a static model of single-family residential development. In the model, landowners decide whether to undertake new construction or redevelopment and, conditional on developing, choose among single-family housing prototypes constructed by clustering observed building dimensions within each municipality. These choices account for minimum lot size regulation, expected sale prices of newly constructed homes, the value of any existing structure, construction costs, and municipality-specific implicit development costs. Landowners choose the option that maximizes profit per unit of land, including the outside option of maintaining the status quo. The key structural parameters are the municipality-level implicit development costs, which I decompose into a lump-sum component that applies to all development and an additional penalty for building on lots smaller than the minimum lot size. Intuitively, these costs are identified from the gap between predicted development profitability and observed development activity. When a municipality has high home prices relative to construction costs yet development remains uncommon, the model attributes the shortfall to higher implicit costs. Similarly, when development on nonconforming lots is relatively common, all else equal, the estimated cost of building below the minimum lot size is lower.

I estimate the model separately for each of the 98 municipalities in Greater Boston by maximum likelihood using data from 2010–2019. The estimates imply substantial cross-municipality variation in implicit development costs. The median lump-sum cost is approximately \$117,000 per project, with an interquartile range of roughly \$30,000 to \$250,000, and the median additional cost associated with development that does not conform to the zoning code is about \$45,000.

These estimates should be interpreted as capturing implicit costs beyond the explicit permitting fees and physical costs already embedded in observed construction estimates, including delay, uncertainty, financing, and other costs of navigating the development process.

I use the estimated model to examine the housing supply effects of relaxing each constraint, separately and in combination, first under fixed prices and then allowing for equilibrium price adjustment. Holding prices fixed, lowering construction costs has a large effect on supply. Applying 2024 cost estimates, which are 16% higher than the 2014–2019 baseline, reduces predicted production by 47%, while a 15% cost reduction, of the magnitude attributed to modular or prefabricated construction, increases it by 79%. Reducing the per-project development cost, for instance by streamlining permitting and discretionary review, raises supply by 79% at $50,000 and by 189% at $100,000, and removing minimum lot size requirements alone raises supply by 27%. Combining reforms is more than additive. Removing minimum lot size requirements together with a $100,000 cost reduction, for example, increases supply by 271%, roughly 25% more than the two reforms deliver separately, because relaxing the lot-size constraint draws out additional small-lot development once the cost of building is also reduced.

I then allow for equilibrium price responses to account for the fact that greater housing supply lowers local house prices. Using alternative values of the price response parameter, which governs how prices respond to changes in the housing stock, I find that equilibrium supply responses are moderately attenuated relative to the fixed-price results. At the baseline level of price responses, for example, the $50,000 cost reduction increases housing production by 73% and the $100,000 reduction by 172%, compared with 79% and 189% under exogenous prices, while the combined reform increases production by 252%. The price effects follow the same pattern as the quantities. Individually the reforms move local prices less than 1%, but the comprehensive combined reform (remove minimum lot size plus a $100,000 cost reduction) lowers prices by roughly 5–8% across the range of price responses considered, more than the separate reforms imply. This complementarity, in both quantities and prices, means that the largest affordability gains arise when multiple supply constraints are relaxed together rather than any one separately.

The counterfactual results also reveal important composition effects. Reducing the per-project development cost shifts new construction toward smaller lots and denser development. Under exogenous prices, for example, the mean lot size among new developments falls from 40,100 to 27,800 square feet with a $50,000 reduction, and to 21,900 square feet with a $100,000 reduction. These composition effects remain similar when prices adjust in equilibrium. Overall, the results suggest that relaxing these barriers would not only increase the quantity of new single-family housing but also shift it toward more compact, denser building types.

Related Literature. This paper contributes to the literature studying the economics of housing production and redevelopment decisions by landowners and developers. Several studies in this literature (e.g., Ahlfeldt and McMillen, 2014; Combes et al., 2021; He, 2024) model developers' choices over floorspace construction. Murphy (2018) highlights the role of forward-looking behavior, emphasizing how landowners time development decisions based on expectations about future prices and construction costs. Mayer and Somerville (2000) and Gabriel and Kung (2025) study the role of land markets and the development process in shaping housing outcomes. I add to this literature by introducing a discrete-choice model of single-family housing development, explicitly accounting for construction cost variation within a metropolitan area across locations and building types, and examining how these costs interact with other supply constraints and market conditions.

Most closely related to this paper, Rollet (2025) and Ma (2024) estimate models of housing development using parcel-level data from New York City and Atlanta, respectively, with an emphasis on zoning as a binding constraint shaping redevelopment patterns. Rollet (2025) focuses on floorspace supply through redevelopment decisions, while Ma (2024) focuses on single-family development on vacant parcels. I contribute to this literature by studying single-family construction on both vacant and improved parcels within a discrete-choice framework that separates explicit zoning from an implicit development cost wedge, which allows me to identify how these factors interact.

More broadly, this paper relates to a literature that measures the elasticity of housing supply across U.S. metropolitan areas. Saiz (2010) focuses on exogenous geographic constraints, such as steep terrain and bodies of water, that limit land availability. Baum-Snow and Han (2024) study the micro-geography of housing supply, documenting how supply elasticity varies at the tract level. Green et al. (2005) estimate supply elasticities that vary across metros and review their sources. Paciorek (2013) decomposes metro-level supply elasticity into regulatory and geographic components, documenting their complementarity. I complement this literature by using parcel-level data and zoning-district variation within a single metropolitan area, which allows me to explicitly account for multiple supply factors at a fine spatial scale.

This paper also contributes to a large literature studying the role of land-use and zoning regulations in shaping housing supply. Gyourko and Molloy (2015) review evidence that regulation raises house prices, reduces construction, and lowers housing supply elasticities, while emphasizing the measurement challenges and welfare tradeoffs involved in evaluating regulation. Glaeser and Ward (2009) examine land-use regulations in the Greater Boston area. This paper is most closely related to Kulka et al. (2026), who evaluate the effects of multiple zoning regulations and their interactions in the MAPC region. I build on these insights by combining district-level zoning regulations with parcel-level measures of developable land and construction costs, and by using a unified structural framework to quantify how zoning restrictions affect both the level and the composition of single-family housing supply and, in turn, prices.

Finally, this paper relates to a literature on implicit regulatory and administrative barriers to housing supply beyond formal zoning. For example, the Wharton Residential Land Use Regulatory Index (Gyourko et al., 2008; Gyourko and Krimmel, 2021) measures multiple dimensions of land-use regulation, including permitting and approval processes in addition to written zoning rules. Prior work shows that permitting frictions can impose large effective costs and limit housing production. Mayer and Somerville (2000), Wrenn and Irwin (2015), and Gabriel and

Kung (2025) document how approval durations and delay affect housing production, and Soltas and Gruber (2026) quantify the value of permitting using a market for "ready-to-issue" permits in Los Angeles. The key intuition of this paper follows Glaeser and Gyourko (2018), attributing price–cost wedge to regulatory taxes. By estimating municipality-level implicit development costs that rationalize observed (re)development choices and price–cost wedges, my framework complements this literature and quantifies how these administrative frictions affect both the level and the composition of single-family housing supply.

2 Background and Data

This section describes the institutional setting of Greater Boston and the datasets used in the analysis. I first discuss how Massachusetts delegates land-use authority to individual municipalities, which produces the fragmented zoning environment and cross-municipality variation in the development process that the paper exploits. I then introduce the three main data sources, namely the MAPC Zoning Atlas for district-level zoning regulations, ATTOM Assessor History and Recorder data for parcel characteristics and transactions, and Craftsman Costbooks for construction cost estimates. Finally, I provide summary statistics for the analysis sample.

2.1 Institutional Setting

Massachusetts grants municipalities broad authority over land-use regulation through the Zoning Act (M.G.L. Chapter 40A) and the 1966 Home Rule Amendment to the state constitution. These provisions allow each city and town to adopt and enforce its own zoning code, which governs land use, development density, and dimensional standards such as lot size, floor-area ratio, building height, and setbacks. While local zoning must be consistent with state law, municipalities retain substantial discretion.[1]

1. During the sample period (2010–2019) most zoning changes required a two-thirds vote by the local legislative body. The 2020 Housing Choice legislation (An Act Enabling Partnerships for Growth, St. 2020, c. 358) subsequently

Each municipality's zoning code defines a set of zoning districts, along with a schedule of permitted uses and dimensional requirements. Projects that satisfy these requirements typically proceed by right, subject to administrative review. Nonconforming projects generally require a variance, and many municipalities also require special permits for certain uses, introducing discretionary review, public hearings, and procedural delays. These processes can raise the effective cost of development by increasing delay, uncertainty, and the legal and consulting expenses required to obtain project approval.

The Metropolitan Area Planning Council (MAPC) region, which forms the study area for this paper, includes 101 municipalities in the Greater Boston area, each operating under its own zoning regime. MAPC is the regional planning agency serving these communities. Its role is not to write local zoning codes, which remain under municipal control, but to support regional planning, coordination, research, and data standardization across municipalities. This institutional fragmentation creates substantial cross-municipality variation in formal zoning rules and in the local development process.

Codified zoning rules do not fully capture the constraints developers face in practice. Municipalities with similar written codes can differ substantially in how they administer permitting and review, owing to differences in administrative capacity, local politics, and resident opposition to development.[2] Consequently, the effective regulatory environment may differ from the one implied by the zoning code alone. These features, namely fragmented local governance and the gap between codified rules and the effective regulatory environment, are common across U.S. metropolitan areas, making Greater Boston a useful setting for studying mechanisms that extend beyond this particular market.

---

lowered this threshold to a simple majority for certain housing-enabling amendments, including as-of-right multifamily and mixed-use development in eligible locations, accessory dwelling units, and open-space residential development.

2. For example, some municipalities, such as Cambridge, maintain professional planning departments with dedicated development review staff and standardized procedures, while others, particularly smaller towns such as Boxford, rely on elected volunteer boards with more limited administrative capacity.

2.2 MAPC Zoning Atlas

My first main data source is the MAPC Zoning Atlas, which provides district-level zoning regulations. The Atlas provides GIS polygons for 1,775 zoning districts and associated district-level attributes drawn from local zoning regulations across 101 municipalities. For each district, it reports standardized use restrictions and dimensional requirements that govern development density, including minimum lot size, maximum lot coverage, maximum floor-area ratio, maximum building height, and related constraints that vary by municipality and district. I treat the MAPC Zoning Atlas snapshot as of May 2025 as time-invariant over the analysis period, 2010–2019.

I assign zoning regulations to parcels by spatially joining zoning district polygons to parcel centroids. Each parcel inherits the zoning regulations of its assigned district. The key input for the single-family development model is the minimum lot size requirement, which I interpret as the primary binding density constraint for single-family construction in this setting because it directly limits the number of homes per unit of land and exhibits the clearest bunching at the regulated threshold. However, other dimensional restrictions, such as floor-area ratio, maximum height, lot coverage, and density, tend to move together with the minimum lot size across districts (Appendix Figure A.3), so the minimum lot size captures much of the broader variation in zoning stringency. I therefore use it as the single explicit zoning constraint in the model and verify in Section 4.3 that controlling for the other dimensions leaves the estimates largely unchanged.

2.3 ATTOM Assessor History and Recorder Data

My second main data source is ATTOM Assessor History and Recorder data, which provide parcel characteristics and transaction records. These data compile assessor and recorder office records at the parcel and transaction level. The Assessor History data provide parcel-level assessment and structural attributes, including geocoded location, year built, lot size, building square footage, and number of rooms. The Recorder data provide transaction-level information, including sale date and sale price.

I restrict attention to single-family residential parcels and developable vacant parcels over 2010–2019, excluding observations with missing or implausible characteristics (for example, extreme lot sizes or zero building area for parcels recorded as improved). I use ATTOM for three main purposes. First, I construct the baseline stock of housing and developable vacant land to capture pre-period lot configurations and existing structure characteristics. Second, I measure development outcomes from changes in assessor attributes over the analysis window, mapping updates to year built and improvement status into model outcomes: no change, new construction on vacant land, or redevelopment of an improved parcel. Third, I estimate hedonic price models to impute market values for non-transacting parcels and to construct price indexes for each building type, separately for newly built and existing homes as well as for vacant land. Section 4.2 describes the sample restrictions, model input construction, and hedonic imputation procedure in detail.

2.4 Craftsman Costbooks

My third main data source is Craftsman Costbooks, which provide construction cost estimates that vary by structure type, building dimensions, finish grade, and location. I map these estimates to the single-family building types used in the model and construct municipality-specific cost measures that reflect systematic differences in labor, material, and equipment costs across locations.

Appendix Figures A.1 and A.2 show example cost estimates and descriptions from the Costbooks. The cost estimates include core hard costs and selected indirect cost items that scale with building size and configuration while excluding land acquisition costs. They also do not capture uncertainty and implicit costs in the permitting and construction process. As a result, the development cost parameters estimated in Section 4.2 capture additional barriers to development beyond the direct cost components embedded in these estimates.

I collect new-construction cost estimates for up to 14 prototype single-family homes per municipality, spanning the distribution of building sizes and configurations observed in the data, using Costbooks vintages for 2014, 2019, and 2024. The construction of these housing prototypes and the clustering procedure are described in Section 4.2. I use the average of the 2014 and 2019 costs for each housing type and location as the baseline construction cost measure, and I use the 2024 vintage in the counterfactual analysis as a high-cost benchmark.

2.5 Analysis Sample and Summary Statistics

The analysis sample comprises 501,959 single-family residential and developable vacant parcels across 98 municipalities in the MAPC region.[3] Of these, 9,061 parcels (1.8%) are observed to develop a new single-family home during 2010–2019. Below I summarize zoning stringency, construction costs, and parcel characteristics across the sample.

Table 1 summarizes parcel-assigned dimensional zoning requirements in the analysis sample. Among single-family homes (Panel A), the mean minimum lot size requirement is 26,664 square feet, with substantial variation. The 10th percentile is 7,500 square feet and the 90th percentile is 45,000 square feet. Among vacant parcels (Panel B), minimum lot size requirements are higher on average (32,710 square feet) and more dispersed. Floor-area-ratio (FAR) limits frequently do not exist. Only 97,415 of 468,088 single-family parcels (21%) and 9,055 of 33,871 vacant parcels (27%) have a non-missing maximum FAR in the Atlas.

Panel C reports bindingness proxies among new construction during 2010–2019. About 13% of newly built homes have lot sizes within five percent of the minimum lot size threshold, and this share is similar (13.1%) when conditioning on non-missing MLS. FAR constraints bind much less frequently. Only 0.2% overall and 1.0% conditional on non-missing FAR. These patterns suggest

3. Three municipalities from the MAPC Zoning Atlas are excluded from the analysis sample due to insufficient parcel counts or poor parcel data quality.

that minimum lot size is the primary binding density constraint for single-family construction in this market, broadly consistent with prior evidence in Song (2025) and Kulka et al. (2026).

Table 2 summarizes Craftsman Costbooks construction cost estimates per square foot by building type, municipality, and costbook vintage. Panel A reports mean costs for three building-size bins based on prototype square footage. Costs per square foot are systematically higher for smaller homes, consistent with fixed-cost components and less favorable scale economies for compact structures. Across all prototypes, mean costs are similar in the 2014 and 2019 vintages but rise sharply in 2024, increasing from about \$137 to about \$159 per square foot (a 16% increase).

Panel B documents dispersion in costs across locations and building types. The interquartile range of municipality-level mean costs widens in 2024, and the standard deviation across municipalities is roughly \$19 per square foot in 2014–2019 and about \$24 per square foot in 2024. There is also meaningful within-municipality dispersion across prototypes, with an average standard deviation of about \$12–\$13 per square foot. Overall, construction costs exhibit substantial heterogeneity across both locations and building types.

Table 3 reports summary statistics for the analysis sample. Panel A summarizes baseline parcel characteristics as of 2009, the year before the start of the development window in my analysis. About 6.7% of parcels are vacant at baseline. The median lot size is 12,000 square feet, the median building area among improved parcels is 2,832 square feet, and the mean year built is 1946. Panel B reports development rates over 2010–2019. The overall development rate is 1.8%, with substantially higher rates on vacant parcels (10.5%) than on improved parcels (1.2%).

Panels C–E report transaction prices from ATTOM deed records over 2010–2019. A transaction is classified as new construction if the home was developed during the analysis window and as an existing-home sale otherwise. New construction shows higher prices than existing homes, with a median sale price of \$570,000 for new single-family transactions versus \$415,000 for existing-home transactions. The median unit price is \$223 per building square foot for new homes and $245 for existing homes. Vacant land transactions are relatively sparse (2,080 sales), with a median sale price of $300,000 and a median unit price of $10.1 per lot square foot.

3 Empirical Motivation

This section documents empirical patterns in the development of new single-family housing across Greater Boston municipalities. I establish three facts. First, development is a rare event whose likelihood depends heavily on existing parcel characteristics. Second, minimum lot size is the primary binding zoning constraint, while development profitability varies sharply with lot size. Third, prices, costs, land availability, and zoning stringency are correlated in ways that complicate the interpretation of any single factor on its own.

3.1 Development is rare and depends on existing parcel characteristics

New single-family construction is an infrequent event. Over 2010–2019, only 1.8% of parcels in the analysis sample are observed to develop. Figure 1 shows the distribution of municipality-level development rates, which range from less than 0.5% to over 7%, with a long right tail driven by a handful of rapidly developing suburban communities.

Development rates differ substantially by parcel status. As shown in Figure 2a, vacant parcels develop at a rate of 10.5%, compared to just 1.2% for improved parcels, a nearly ninefold difference. This gap reflects both the additional cost of demolishing an existing structure before rebuilding and the large opportunity cost of forgoing the value of the current home. That is, to redevelop an existing single-family home, the profit margin from new construction must exceed the combined value of the existing structure and the cost of demolition.

Among improved parcels, redevelopment rates vary systematically with the age of the existing structure. Figure 2b plots redevelopment rates by decade of original construction. Homes built in the 1940s–1950s are redeveloped at the highest rates (roughly 2%), while those built after 1980 have redevelopment rates below 0.2%. This pattern is consistent with depreciation lowering the opportunity cost of demolishing older structures. The existing home value likely declines with building age, making redevelopment relatively more attractive. The lower redevelopment rates among the oldest homes (pre-1920) may reflect that these structures are disproportionately located in dense urban cores with more development hurdles, or that they face additional constraints such as historical preservation regulations.

3.2 Zoning constrains lot size choices, and profitability varies with lot size

Minimum lot size (MLS) requirements are the most visibly binding zoning constraint for single-family development, as discussed in Section 2.5. Figure 3a visualizes this by plotting the distribution of the ratio of lot size to MLS among newly developed parcels. There is sharp bunching just at the threshold, indicating that developers build on the smallest conforming lot available. At the same time, 61.5% of new construction occurs on lots *smaller* than the assigned MLS, with a median ratio of 0.79, likely reflecting pre-existing nonconforming lots that predate current zoning or variance approvals. Vacant parcels show the same pattern, with a large share lying below their assigned minimum lot size (Appendix Figure A.5). While some vacant parcels are conforming and could be developed at or above the threshold, much of the developable land already sits on undersized parcels.

At the same time, development profitability varies strongly with lot size. Figure 3b plots mean unit profit, defined as sale price minus physical construction cost divided by lot size (dollars per square foot of land), against lot size across all municipality-type cells. The relationship is steeply declining: small-lot building types yield a median unit profit of $26 per square foot of land, compared to $10 for medium-lot types and $4 for large-lot types. This inverse relationship arises because, although compact lots are associated with both higher unit revenue and higher construction costs, the revenue effect dominates (see Appendix Figure A.4). As a result, minimum lot size requirements can deter precisely the building types that would be most profitable absent dimensional constraints.

3.3 Prices, costs, land availability, and zoning jointly vary across locations

The profitability of new construction also varies substantially across municipalities, driven primarily by differences in house prices. Prices vary over three times as much as construction costs across municipalities (standard deviation of $342,000 vs. $104,000), even though the two are strongly positively correlated (correlation +0.73), as shown in Table 4. Therefore, profit margins are still highest in high-price municipalities, much like the inverse relationship between profit margins and lot size discussed above.

However, high-profit municipalities do not exhibit the highest development rates. High-price municipalities tend to have less vacant land available for development (correlation between price and vacant share is -0.27), reflecting that the most profitable locations have already been substantially built out. They are also not systematically associated with looser zoning (correlation between MLS and price is -0.03), meaning that higher-price municipalities, including those in the urban core, do not necessarily have smaller minimum lot sizes. As a result, the raw correlation between profit margins and development rates is actually negative (-0.19). It illustrates the extent to which land scarcity and regulations may offset the incentives created by high price margins.

Together, these correlations show that land availability, prices, construction costs, and zoning stringency interact in ways that make any single factor difficult to interpret on its own. Variables that appear to predict more development in simple comparisons, such as larger minimum lot sizes or higher construction costs, are correlated with other features of municipalities that work in the opposite direction. This motivates an approach that jointly accounts for prices, costs, existing development patterns, and zoning when modeling development outcomes, with the goal of disentangling the roles of each factor in shaping housing supply.

4 A Model of Single-Family Residential Development

I develop a discrete-choice model of single-family residential development in which landowners choose whether and what to build on each parcel. The model incorporates various supply-side factors, including house prices, construction costs, existing structure values, and zoning constraints and estimates municipality-specific implicit development cost parameters that rationalize the observed pattern of development choices. This section presents the model setup (Section 4.1), describes the quantification of model inputs and the estimation procedure (Section 4.2), reports estimation results (Section 4.3), and conducts counterfactual analyses (Section 4.4).

4.1 Setup

A landowner with parcel i chooses between not developing and developing a new single-family home of type k, subject to zoning regulations, construction costs, and implicit development costs that vary by municipality $m(i)$.

Let $k \in \mathcal{K}_{m(i)}$ index a finite set of single-family building types available in municipality $m(i)$. The payoff from choosing type k is

$$\Pi_{ik} = \frac{P_{ik} - C_{ik} - \omega_{m(i),1} - \omega_{m(i),2} \cdot \mathbb{1}(x_{1,k} < \bar{x}_{1,i}) - \gamma_0 \cdot s_i \cdot \mathbb{1}(\text{improved}_i)}{x_{1,k}} + \varepsilon_{ik}, \qquad (1)$$

where P_{ik} is the market price of a home of type k on parcel i, C_{ik} is the corresponding (physical) construction cost, and $x_{1,k}$ is the lot size associated with type k. The parameters $\omega_{m,1}$ and $\omega_{m,2}$ are municipality-specific lump-sum development costs denominated in dollars. $\omega_{m,1}$ captures a baseline implicit cost of development, and $\omega_{m,2}$ captures an additional cost incurred when the chosen lot size falls below the minimum lot size requirement $\bar{x}_{1,i}$ applicable to parcel i. The model includes the minimum lot size as the single explicit zoning constraint, both because it is the most visibly binding requirement in this setting (Section 2.5) and because the other dimensional rules move closely with it across districts, so it proxies for broader zoning stringency. Appendix D confirms that adding the remaining dimensional requirements as controls leaves the estimated costs largely unchanged.

The term $\gamma_0 \cdot s_i \cdot \mathbb{1}(\text{improved}_i)$ captures the demolition and site clearance cost of redeveloping an improved parcel, where s_i is the square footage of the existing building and γ_0 is a per-square-foot demolition cost. Dividing by lot size $x_{1,k}$ converts the total development payoff into a return per square foot of land. Define the systematic component of the development payoff as

$$V_{ik} = \frac{P_{ik} - C_{ik} - \omega_{m(i),1} - \omega_{m(i),2} \cdot \mathbb{1}(x_{1,k} < \bar{x}_{1,i}) - \gamma_0 \cdot s_i \cdot \mathbb{1}(\text{improved}_i)}{x_{1,k}}. \tag{2}$$

The outside option is to not develop:

$$\Pi_{i0} = \bar{\pi}_i + \varepsilon_{i0}, \tag{3}$$

where $\bar{\pi}_i$ denotes the status quo value of parcel i per square foot of land. For improved parcels, $\bar{\pi}_i$ is the hedonic-predicted market value of the existing structure divided by lot area. For vacant parcels, $\bar{\pi}_i$ is the estimated land value per square foot. In both cases, the outside option is denominated in the same per-square-foot units as the development payoff.

I assume $\{\varepsilon_{i0}, \varepsilon_{ik}\}_{k \in \mathcal{K}_{m(i)}}$ are i.i.d. type-I extreme value. The probability that parcel i is developed with type k is then

$$Pr_i(k) = \frac{\exp(V_{ik})}{\exp(\bar{\pi}_i) + \sum_{k' \in \mathcal{K}_{m(i)}} \exp(V_{ik'})}, \qquad k \in \mathcal{K}_{m(i)}. \tag{4}$$

The probability of (re-)development is

$$Pr_i(\text{develop} = 1) = \sum_{k \in \mathcal{K}_{m(i)}} Pr_i(k) = \frac{\sum_{k \in \mathcal{K}_{m(i)}} \exp(V_{ik})}{\exp(\bar{\pi}_i) + \sum_{k \in \mathcal{K}_{m(i)}} \exp(V_{ik})}, \tag{5}$$

and the conditional probability of choosing type k given development is

$$Pr_i(k \mid \text{develop} = 1) = \frac{\exp(V_{ik})}{\sum_{k' \in \mathcal{K}_{m(i)}} \exp(V_{ik'})}. \tag{6}$$

Because the implicit development cost parameters are divided by the type-specific lot size $x_{1,k}$, neither cancels from the conditional choice probability in (6), so both affect the extensive margin (whether to develop) and the intensive margin (which type to build). However, $\omega_{m,2}$ additionally penalizes types whose lot size falls below the minimum lot size requirement. This provides the basis for separate identification. $\omega_{m,1}$ is pinned down by the overall development rate and the lot-size distribution of chosen types across the full choice set, while $\omega_{m,2}$ is pinned down by the additional tendency to avoid below-minimum-lot types conditional on the lot-size gradient already imposed by $\omega_{m,1}$.

The model abstracts from several real-world complexities. First, it allows flexible subdivision and assembly of parcels, which permits the chosen lot size $x_{1,k}$ to differ from the pre-period parcel configuration. I examine the extent to which the counterfactuals rely on assembly in Section 4.4. Second, zoning regulations and the parameters $\omega_{m,1}$ and $\omega_{m,2}$ are treated as time-invariant over 2010–2019, with zoning measured from the MAPC snapshot as of May 2025. Third, I collapse the ten-year analysis period into a single static decision over 2010–2019, treating the development choice as a one-shot decision over the decade rather than dynamic, forward-looking timing. Dynamic models of development emphasize the option value of delaying construction in anticipation of higher future prices (Murphy, 2018), but optimal waiting is unlikely to account for the persistently low development observed here over the ten-year time horizon. To the extent that forward-looking behavior and unmodeled heterogeneity do shape development timing, they are absorbed into the estimated development costs, consistent with the interpretation of $\omega_{m,1}$ as a residual development cost rather than a purely regulatory parameter.

4.2 Quantification of the model

The structural parameters are municipality-specific implicit development costs $\{\omega_{m,1}, \omega_{m,2}\}_m$, which I estimate by maximum likelihood separately for each municipality using the choice probabilities in (4)–(6). Let I_0 denote parcels that are not (re-)developed during 2010–2019, and let I_1 denote parcels that are (re-)developed. For each $i \in I_1$, let $\hat{k}_i$ denote the observed building type. The log-likelihood for municipality m is

$$\mathcal{L}_m(\omega_{m,1}, \omega_{m,2}) = \sum_{\substack{i \in I_0 \\ m(i)=m}} \log Pr_i(0) + \sum_{\substack{i \in I_1 \\ m(i)=m}} \log Pr_i(\hat{k}_i), \tag{7}$$

where

$$Pr_i(0) = \frac{\exp(\bar{\pi}_i)}{\exp(\bar{\pi}_i) + \sum_{k \in \mathcal{K}_{m(i)}} \exp(V_{ik})}.$$

I maximize (7) by the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm with analytic gradients, using the observed Hessian to compute standard errors.

The number of estimated parameters per municipality depends on identification. The parameter $\omega_{m,1}$ is estimated for all municipalities with at least one observed development. The parameter $\omega_{m,2}$ is estimated only when the municipality's choice set includes types both above and below the minimum lot size requirement and the observed developments are not exclusively of one kind, which would make $\omega_{m,1}$ and $\omega_{m,2}$ collinear. When $\omega_{m,2}$ is not identified, I set it to zero and estimate only $\omega_{m,1}$. The identification argument follows from the discussion in Section 4.1. A higher $\omega_{m,1}$ reduces overall development propensity and shifts the conditional choice toward larger lots. A higher $\omega_{m,2}$ further decreases the share of below-minimum-lot types conditional on the lot-size gradient already imposed by $\omega_{m,1}$.

Below, I describe how I construct the key model inputs: the unit profit $\pi_{ik} = (P_{ik} - C_{ik})/x_{1,k}$, the status quo value $\bar{\pi}_i$, and the minimum lot size $\bar{x}_{1,i}$.

Building types (k)

I discretize single-family building types using the k-prototypes clustering algorithm, which accommodates both numerical and categorical variables. Clustering is conducted separately within each municipality using single-family parcels built since 1987, merged to MAPC zoning data, and restricted to observations with complete information on the relevant characteristics, with an exception that if finished or unfinished basement area is missing, I consider zero. Numerical variables include building square footage, lot size, finished basement area, and unfinished basement area. Categorical variables include the number of bedrooms, bathrooms, and stories. I set the number of clusters to fourteen based on the elbow method, as shown in Figure A.6. Each cluster defines a prototype k with associated attributes, including lot size $x_{1,k}$. Building types whose lot size falls below the minimum lot size $\bar{x}_{1,i}$ applicable to a parcel's zoning district, measured from the MAPC zoning atlas, are flagged as below-minimum-lot alternatives in the choice set. After clustering, I remove building types with zero observed development share during the analysis period within each municipality, so that the choice set $\mathcal{K}_{m(i)}$ contains only types that have been chosen at least once.

Previous parcel characteristics and observed construction choices

I construct the inventory of single-family parcels as of 2009 using historical ATTOM assessor records to measure pre-period parcel characteristics and to define the status quo outcome. I classify each parcel as improved (an existing structure in place as of 2009) or vacant. Development outcomes over 2010–2019 are defined using year-built information: a parcel is classified as developed if its recorded construction year in the 2020 tax record exceeds 2009. Parcels with only an updated effective year built but no change in the actual year built are classified as undeveloped, as these likely reflect renovations rather than new construction, and the model's construction cost estimates are calibrated to new builds. For developed parcels, the observed building type $\hat{k}_i$ is assigned by matching the 2020 parcel characteristics to the nearest cluster prototype.

Value of new construction (π_{ik})

To construct building-type-by-location price indexes for new construction, I restrict the sample to transactions of newly built single-family homes occurring within two years of the recorded construction year, using data from 2010 through 2019. I winsorize transaction prices at the 1st and 99th percentiles.

I estimate a generalized additive model (GAM) for log transaction prices with a two-dimensional thin-plate spline over zoning-district centroid coordinates to capture location premia, year-by-municipality fixed effects for time-varying market conditions, and penalized splines for continuous housing characteristics:

$$\begin{aligned}\log(\text{Price}_{it}) = \alpha &+ f_1(\text{building area}_{it}) + f_2(\text{lot area}_{it}) + f_3(\text{fin. basement}_{it}) + f_4(\text{unfin. basement}_{it}) \\ &+ \beta_1 \cdot \text{bath}_i + \beta_2 \cdot \text{bed}_i + \beta_3 \cdot \text{story}_i + \delta_{\text{muni}(i),t} + g(\text{lon}_i, \text{lat}_i) + \mu_{\text{month}(t)} + \varepsilon_{it},\end{aligned} \tag{8}$$

where $f_1(\cdot), \dots, f_4(\cdot)$ and $g(\cdot,\cdot)$ are penalized spline functions estimated by fast restricted maximum likelihood, and $(\text{lon}_i, \text{lat}_i)$ denotes the centroid of parcel i's zoning district. I predict prices for each zoning district–municipality–building type cell (including districts with no observed new-construction transactions) and average the predicted prices across 2010–2019 to obtain P_{ik}, which varies by location and building type but not by year. Averaging across years implicitly assumes that landowners form price expectations based on a long-run average value of new construction in their location and building-type cell.

Appendix Section B.2 reports model fit diagnostics and compares the baseline specification with OLS log-linear hedonic models and alternative controls that absorb zoning-district fixed effects rather than using a continuous spatial surface. As a robustness check, I also construct P_{ik} using an OLS hedonic regression in place of the GAM specification.

I measure construction costs C_{ik} using Craftsman Costbooks, mapped to the building prototypes. These estimates capture materials and labor costs that vary with building characteristics and location, and they exclude land acquisition and demolition costs. I average the 2014 and 2019 cost editions to obtain a cost measure aligned with the estimation period. I then deflate the estimates by 11% to remove the contractor markup component reported in the Costbooks, so that C_{ik} reflects direct hard costs only. Any remaining markup, soft costs, or other frictions are absorbed into the estimated development cost parameters. The resulting unit profit is $\pi_{ik} = (P_{ik} - C_{ik})/x_{1,k}$, denominated in dollars per square foot of land.

The demolition cost parameter γ_0 is calibrated from industry cost guides for Greater Boston residential demolition, which report per-square-foot costs in the range of \$3–\$17.[4] I set $\gamma_0 = \$7$ per square foot as the baseline and examine sensitivity at $\gamma_0 \in \{\$4, \$10\}$.

Status quo value ($\bar{\pi}_{it}$)

The status quo value $\bar{\pi}_i$ is the per-square-foot opportunity cost of the parcel in its current use. For improved parcels, I impute $\bar{\pi}_i$ using a GAM-spatial hedonic model estimated on existing single-family transactions (buildings at least two years old at time of sale). The specification mirrors Equation (8) but adds building age as a penalized smooth predictor, allowing for a flexible depreciation curve. I predict status quo values for all pre-2009 improved parcels at each year from 2010 through 2019, so that the depreciation curve and year×municipality effects adjust the valuation over time, and average the predictions across years. The hedonic-predicted total home value is converted to a per-square-foot-of-land measure by dividing by the parcel's lot area.[5] Appendix Section B.3 reports model fit diagnostics and compares the baseline GAM-spatial predictions with OLS and zoning-district fixed-effect specifications.

4. For example, HomeGuide reports \$4–\$10; Fixr reports \$5–\$10 with a low of \$3 for homes without foundations; and Angi/HomeAdvisor reports \$4–\$17, with the upper end reflecting urban sites and multi-story structures. Boston-specific estimates from HomeBlue imply approximately \$6–\$10 per square foot for a typical 1,500–2,000 square foot home. All sources accessed March 2026.

5. Missing lot area and building area are imputed using zoning-district-level medians from the non-missing parcels in the same district. Approximately 12% of parcels require imputation; the resulting predictions are flagged and their distribution is compared to the non-imputed sample to verify that the imputation does not introduce systematic bias. Appendix Section B.3 reports imputation diagnostics.

For vacant parcels, I measure $\bar{\pi}_i$ using land values from a GAM model estimated on vacant land transactions. I estimate log transaction price per square foot on a two-dimensional spatial smooth in zoning-district centroid coordinates, a smooth in lot size, and municipality and year fixed effects. This measure covers all zoning districts through spatial interpolation where transaction data are sparse. I average the predictions across years to obtain a single per-square-foot land value per zoning district. Appendix C compares these GAM-based land value estimates with the Federal Housing Finance Agency (FHFA) experimental land price dataset (Davis et al., 2021), which estimates the value of land underlying single-family structures from cost-approach appraisals at the census tract level.

In one robustness check, instead of averaging predicted status quo values across years, I use the minimum predicted status quo value over 2010–2019 as $\bar{\pi}_i$. This scenario corresponds to a development-friendly benchmark in which landowners evaluate redevelopment against the lowest anticipated value of the existing structure, which lowers the outside option and, holding observed development choices fixed, leads to larger estimated development costs.

4.3 Estimation Results

Implicit development cost estimates

Table 5 summarizes the estimated implicit cost parameters across municipalities. Because the Craftsman construction cost estimates C_{ik} already include explicit permit fees, $\omega_{m,1}$ and $\omega_{m,2}$ are best understood as a residual development cost, the gap between the measured profitability of construction and the rate at which development actually occurs. This residual captures the implicit costs of navigating the regulatory process, including delays from multi-stage review, uncertainty over approval outcomes, legal and consulting expenses, and community opposition. It also absorbs non-regulatory frictions that the model does not measure separately, including financing and carrying costs, developer risk and required returns, soft costs, the option value of delaying construction, and any measurement error in construction costs or prices. I therefore interpret $\omega_{m,1}$ as an upper bound on the purely regulatory burden rather than a direct measure of it.

The baseline development cost $\omega_{m,1}$ has a median estimate of \$117,285 (mean \$223,365), implying that a developer in the median municipality faces a lump-sum implicit burden well in excess of the direct construction costs. Dividing by the median lot size in the choice set (22,731 square feet), this corresponds to approximately \$5.16 per square foot of land, or roughly 21% of the mean unit profit π_{ik}. The distribution is right-skewed, with a standard deviation of \$446,438 and a maximum of \$3.4 million, reflecting a handful of municipalities with very high estimated development costs.

The estimated $\omega_{m,1}$ is not only dispersed but systematically related to local market conditions. Appendix Figure A.7 plots the estimated $\omega_{m,1}$ against the development profitability gap (maximum unit profit minus mean status quo value) for each municipality. The two are strongly positively correlated ($\rho = 0.87$), so municipalities where development is most profitable relative to the status quo have the highest estimated development costs. This implies that implicit barriers are binding precisely where market incentives to develop are strongest.

Twenty-three municipalities have negative $\omega_{m,1}$ estimates, which may reflect overestimation of construction costs in the Craftsman data as it reflects retail cost estimates. If the cost estimates exceed actual builder costs, the model compensates by assigning a negative development cost to rationalize the observed development rate. Because of these negative estimates, the counterfactual exercises in Section 4.4 reduce $\omega_{m,1}$ by a uniform dollar amount rather than setting it to zero, which would counterintuitively increase implicit costs for municipalities with negative estimates.

Although $\omega_{m,1}$ blends regulatory and non-regulatory frictions, this composite interpretation does not undermine the counterfactual exercise. The central point is that much of the residual is reducible through the same process reforms the counterfactuals represent, whether or not it is regulatory in the narrow sense. By-right or expedited review shortens approval timelines and reduces the uncertainty over outcomes that drives financing, carrying, risk, and soft costs, so streamlining compresses both the regulatory content of $\omega_{m,1}$ and much of its non-regulatory content.

The estimated implicit development costs and the cost reductions I consider in Section 4.4 are also comparable in magnitude to externally documented regulatory costs. The National Association of Home Builders (NAHB) estimates that government regulation adds \$93,870 to the price of an average new single-family home nationally, based on its 2021 NAHB/Wells Fargo surveys, comprising \$41,330 in development-phase costs and \$52,540 in construction-phase costs, both of which reflect regulatory burden alone and exclude the physical materials and labor costs. The \$50,000 and \$100,000 reductions lie within the range that external evidence attributes to documented regulatory sources, so the exercise does not require $\omega_{m,1}$ to be purely regulatory in order to interpret the reduction as the effect of relaxing development frictions. The estimated median $\omega_{m,1}$ of \$117,000 is roughly 25 percent above the NAHB national average, consistent with Greater Boston's more complex permitting environment.

The additional cost $\omega_{m,2}$ of building below the minimum lot size is estimated for 95 municipalities where the choice set includes building types both above and below the minimum lot size and the development data provide sufficient variation to identify both parameters separately. The median estimate is \$44,959 (mean \$58,781), indicating a substantial penalty for developing lots smaller than the regulated minimum, consistent with the expense and uncertainty of obtaining a zoning variance.

Model fit

Figure 4 demonstrates that the model fits, both on a targeted moment and on an untargeted one. Panel (a) plots observed versus predicted municipality-level development rates. The model matches the cross-sectional distribution closely, with a correlation of 0.923 and a mean absolute error of 0.39 percentage points. Panel (b) shows that the predicted lot-size distribution of chosen building types tracks the observed distribution, a moment not directly targeted by the likelihood.

Robustness

Table 6 reports the implicit development cost estimates under alternative specifications. I first assess whether the baseline implicit cost estimates merely absorb explicit zoning rules omitted from the baseline specification, which represents zoning through the minimum lot size alone. To do so, I re-estimate the model with district-level controls for the maximum floor-area ratio, building height, lot coverage, and by-right multifamily zoning. The median $\omega_{m,1}$ moves only from \$117k to \$114k and correlates 0.94 with the baseline across municipalities, and the fit is comparable ($\rho = 0.924$, versus 0.923 at baseline). The implicit cost is thus not an artifact of omitted explicit zoning. Appendix D reports the augmented specification and the estimated control coefficients.[6]

I then re-estimate the model under six alternative specifications that vary the price index (OLS-spatial instead of GAM-spatial, applied to both new construction and status quo), status quo timing (period minimum instead of mean), demolition cost ($\gamma_0 \in \{\$4, \$10\}$), and sample restriction (excluding parcels with imputed lot size). The median $\omega_{m,1}$ ranges from \$82k (OLS prices) to \$201k (period-minimum status quo with low demolition cost), compared to \$117k at baseline. The estimates are most sensitive to the status quo timing: using the period minimum rather than the mean raises the median $\omega_{m,1}$ by roughly \$80k, because a lower outside option makes development appear more profitable and requires a larger implicit cost to match the observed rate. By contrast, the estimates are nearly invariant to the demolition cost calibration (\$116k–\$121k across $\gamma_0 \in \{\$4, \$7, \$10\}$), confirming that the demolition cost is a small component of the overall cost structure. The combination of period-minimum status quo and low demolition cost represents the most development-friendly market assumption, as it minimizes

6. The coefficients ψ of other explicit zoning rules are only weakly identified and are not interpreted individually. Because any municipality-level zoning measure is collinear with municipality-level constant $\omega_{m,1}$, ψ is pinned down solely by within-municipality variation in zoning rules across districts, which is thin. For example, a maximum floor-area ratio is recorded for only about 2.5% of single-family parcels, and a maximum height for nearly all of them, leaving little independent cross-district variation. As a result ψ trades off against the municipality cost estimates along a nearly flat ridge of the likelihood, so I report point estimates without standard errors. The stability of $\omega_{m,1}$ and $\omega_{m,2}$ holds regardless.

the outside option and the cost of redevelopment, and accordingly produces the highest median $\omega_{m,1}$ ($201k), since a larger implicit cost is needed to rationalize the same observed development rate under more favorable market conditions. The correlation between observed and predicted development rates exceeds 0.90 in all specifications.

4.4 Counterfactual Analysis

I use the estimated model to evaluate how changes in the physical and regulatory costs of construction and in zoning constraints would affect the level and composition of new single-family housing supply. I first report fixed-price (partial-equilibrium) counterfactuals that hold local house prices constant, and then extend the analysis with an equilibrium price adjustment in which new supply lowers local prices and feeds back into development profitability. The counterfactual exercises modify implicit development cost parameters and, in the cost regime scenarios, construction costs, while holding status quo values at their estimated levels. Because the model is estimated on observed choices over 2010–2019, these exercises are retrospective in that they quantify how development outcomes in this period would have differed under alternative cost and regulatory environments, holding the underlying distribution of parcels and baseline market conditions fixed.

I consider three types of market and policy changes and their combinations.

- *Construction cost regimes.* The high-cost regime replaces the baseline 2014–2019 average Craftsman estimates with the 2024 Craftsman cost book, which reflects the escalation in materials and labor costs through the early 2020s. The low-cost regime applies a 15% uniform discount to the baseline average construction costs, representing potential efficiency gains from modular construction, prefabrication, or other technological changes.[7]

7. Estimates of hard-cost savings from factory-built housing vary. The Terner Center for Housing Innovation places the range at 10–25% under favorable conditions (Metcalf, 2026, testimony to the California Select Committee on Housing Construction Innovation). McKinsey & Company (2019) estimates 17–24% savings for a four-story affordable housing project, depending on the degree of off-site fabrication. Industry practitioners, however, report

- *Removing minimum lot size requirements.* I set $\omega_{m,2} = 0$ for all municipalities and zero out the below-minimum-lot indicator $\mathbb{1}(x_{1,k} < \bar{x}_{1,i})$, so that all building types in the choice set face the same implicit cost regardless of lot size.

- *Reducing implicit costs.* I reduce $\omega_{m,1}$ by a uniform amount across all municipalities, either \$50,000 (partial reform) or \$100,000 (comprehensive reform). The partial reduction corresponds approximately to eliminating development-phase fees and permitting delays, which the NAHB estimates at \$41,000 per home nationally.[8] The comprehensive reduction corresponds to eliminating nearly all explicit and implicit development costs. In both scenarios, the zoning code remains unchanged.

- *Combined reforms.* I combine the removal of minimum lot size requirements with each level of development cost reduction.

Partial equilibrium with exogenous prices

Table 7 reports the results. The baseline model predicts 14,466 new single-family homes over 2010–2019, equivalent to about 2.9 percent of the housing stock over the decade, or roughly 0.29 percent per year. Construction costs have a large effect on predicted supply. The high-cost regime reduces new units by 47 percent and the low-cost regime increases them by 79 percent, underscoring the sensitivity of development to construction cost assumptions. Among the regulatory reforms, removing minimum lot size requirements alone increases predicted new units by 27 percent, with the additional units concentrated in smaller-lot building types previously penalized by $\omega_{m,2}$. Reducing $\omega_{m,1}$ by \$50,000 increases new units by 79 percent, and a comprehensive reduction of \$100,000 increases them by 189 percent. The largest effect comes

more modest savings of 5–10% on hard costs (Potter, 2025, Construction Physics). Our 15% cost-saving proxy falls within the Terner Center range and is consistent with the McKinsey estimates for partial off-site fabrication.

8. The NAHB estimates that government regulation adds \$93,870 to the price of an average new single-family home (2021), of which \$41,330 is incurred during lot development (impact fees, land dedication, zoning compliance) and \$52,540 during construction (permit fees, building code changes, design mandates, and delay costs). The comprehensive reduction of \$100,000 corresponds approximately to eliminating all documented regulatory costs. The paper's median estimated $\omega_{m,1}$ of \$117,000 is modestly above the NAHB national average, consistent with Greater Boston's more complex permitting environment.

from the comprehensive combined reform, which removes minimum lot size requirements and reduces $\omega_{m,1}$ by \$100,000, increasing predicted new units by 271 percent. Even this most aggressive reform raises construction to about 1.1 percent of the housing stock per year, up from 0.29 percent at baseline. The proportional supply response is therefore large, while the cumulative addition to the stock over the decade remains bounded.

The combined reforms reveal a complementarity between the two regulatory margins. Removing minimum lot size requirements adds about 3,950 units on its own, and reducing $\omega_{m,1}$ by \$100,000 adds about 27,300 units, so their separate effects sum to roughly 31,300 units. Applied together they add about 39,200 units, exceeding the sum of the separate effects by roughly 7,950 units, or 25 percent. The same pattern holds for the smaller reform, where the \$50,000 combination exceeds the sum of its parts by about 32 percent. Relaxing the lot-size constraint draws out additional small-lot development only once the lump-sum development cost is also reduced, so the two reforms act as complements rather than substitutes.

The counterfactual reforms affect not only the quantity of new housing but also its composition. The construction cost regimes illustrate this clearly. Under high costs, the mean lot size among new developments rises from 40.1 to 59.9 thousand square feet, as only the largest and most profitable projects remain viable. Conversely, low construction costs reduce the mean lot size to 29.9 thousand square feet by making smaller-lot development feasible. Removing minimum lot size requirements has a similar compositional effect, shifting the distribution toward smaller lots, with the mean lot size falling to 32.9 thousand square feet. Reducing $\omega_{m,1}$ produces the largest shift, with the mean lot size falling to 21.9 thousand square feet under the comprehensive reduction. This occurs because both construction costs and the lump-sum implicit cost impose a larger per-unit burden on smaller lots, so reducing either disproportionately benefits smaller-lot types and increases density.

Incorporating price adjustments

The fixed-price counterfactuals treat local house prices as exogenous, overstating the supply response if new construction decreases prices. To account for this equilibrium effect, I solve for equilibrium prices consistent with the price response parameter η,

$$P_m^{\mathrm{cf}} = P_m^{\mathrm{base}} \times \left[1 - \eta \cdot \frac{Q_m(P_m^{\mathrm{cf}};\ \boldsymbol{\theta}^{\mathrm{cf}}) - Q_m(P_m^{\mathrm{base}};\ \boldsymbol{\theta}^{\mathrm{base}})}{\mathrm{Stock}_m} \right], \tag{9}$$

where P_m^{base} is the observed price level in each municipality (normalized to 1), $Q_m(\cdot;\ \boldsymbol{\theta})$ is the predicted number of new units as a function of prices and counterfactual parameters $\boldsymbol{\theta}$, Stock_m is the existing housing stock, and η is the price response parameter, the percentage decline in local house prices per 1% increase in the housing stock.

This is a parsimonious approach to incorporating price responses to increased housing supply. Alternatively, estimating an underlying housing demand system (e.g., Anagol et al., 2021; Ma, 2024) can deliver micro-founded price responses that reflect richer substitution patterns and heterogeneous adjustment across neighborhoods. However, such models require additional assumptions about how total population and the preference parameters underlying housing demand evolve in counterfactuals. Instead, I incorporate equilibrium price adjustments using a linear approximation that maps changes in predicted new supply into proportional changes in local house prices within each municipality. Equation (9) can be interpreted as a local approximation around the baseline in which prices respond linearly to changes in the housing stock, scaled by the parameter η. When many households move into the metro area as supply expands, prices adjust relatively little and η is close to zero. When in-migration is more limited, the same supply increase requires larger price declines, implying a larger η.

I examine $\eta = 0.3$ at the baseline, which falls within the range of empirical estimates from prior literature.[9] I additionally evaluate $\eta = 0.6$ as a conservative bound on the supply response

9. Most existing studies estimate the effect of new supply on rents rather than prices. For example, Mense (2025) finds that a 1% increase in new housing supply lowers rents by 0.19% at the municipal level in Germany. Li (2022)

and $\eta = 0.1$ as a liberal bound. For each η and counterfactual scenario, I solve for the fixed point of (9) by iterating the map

$$P_m^{\text{cf},(n+1)} = (1-\lambda)\, P_m^{\text{cf},(n)} + \lambda \left[1 - \eta \cdot \frac{Q_m(P_m^{\text{cf},(n)};\, \boldsymbol{\theta}^{\text{cf}}) - Q_m^{\text{base}}}{\text{Stock}_m}\right],$$

starting from $P_m^{\text{cf},(0)} = 1$, where λ is set to 0.25.

Table 8 reports the price-adjusted results for $\eta \in \{0.1, 0.3, 0.6\}$. Allowing prices to adjust moderately attenuates the supply response relative to the fixed-price counterfactuals. Under the baseline $\eta = 0.3$, the partial $\omega_{m,1}$ reduction yields a 73 percent increase, compared with 79 percent under fixed prices, and the comprehensive reduction yields 172 percent, compared with 189 percent. The low-cost regime yields 71 percent under equilibrium pricing versus 79 percent under fixed prices. Removing minimum lot size requirements gives a result close to the fixed-price case, 29 percent versus 27 percent, because it redistributes development toward smaller lots rather than generating a large net increase in units. The complementarity between the two regulatory margins also survives price adjustment, as the comprehensive combined reform at $\eta = 0.3$ exceeds the sum of its separate parts by about 7,450 units, or 26 percent, close to the 25 percent under fixed prices.

The implied price responses are economically meaningful but still limited. No single reform moves average prices much on its own. Removing minimum lot size requirements leaves the mean price multiplier essentially unchanged at 1.00, and the comprehensive $\omega_{m,1}$ reduction lowers it to about 0.98 at $\eta = 0.3$. The comprehensive combined reform produces the largest decline, with the mean multiplier falling to 0.95 at $\eta = 0.3$ and 0.92 at $\eta = 0.6$, equivalent to average price declines of roughly 5 and 8 percent. These declines are larger than the separate reforms would imply on their own, reflecting the same complementarity seen in the quantity

finds that a 10% increase in the housing stock within 500 feet of new high-rises in New York City lowers nearby rents by roughly 1%, implying $\eta_{\text{rent}} \approx 0.1$ at the hyperlocal level. At the same time, Molloy et al. (2022) show that supply constraints raise purchase prices by roughly twice as much as rents. Together, these estimates imply $\eta \approx 0.2$ to 0.4 for prices.

responses. They represent a substantial improvement in affordability, yet they fall well short of closing the gap in a high-demand metropolitan area, consistent with single-family supply being one lever among several rather than a sufficient one on its own.

The supply responses show meaningful sensitivity to η. The comprehensive $\omega_{m,1}$ reduction ranges from 183 percent at $\eta = 0.1$ to 159 percent at $\eta = 0.6$, and the combined reform ranges from 264 to 235 percent across the same range.

Endogenizing prices also affects the composition of new development, though less dramatically than the quantity effects. Under fixed prices, reducing $\omega_{m,1}$ by $50,000 shifts the mean lot size from 40.1 to 27.8 thousand square feet. Under equilibrium pricing at $\eta = 0.3$, the lot size shift is similar (28.3 thousand square feet), indicating that the composition effect is robust to price adjustment for moderate reforms. For comprehensive reforms, equilibrium lot sizes are slightly larger than under fixed prices (22.5 vs. 21.9 thousand square feet for the $100,000 reduction at $\eta = 0.3$), as the price decline marginally discourages the smallest-lot developments.

Limitations of the model

There are several limitations to keep in mind when interpreting the counterfactual results. First, the counterfactual simulations do not allow for the emergence of new building types outside the constructed choice set and abstract from larger adjustments in design and dimensions. As a result, the counterfactuals may understate substitution toward novel or denser housing configurations under large reforms. Second, the equilibrium price adjustment uses a common parameter η within each municipality, which restricts price responses from varying across housing types. It may therefore mute composition effects that would arise through type-specific or neighborhood-specific price changes.

As mentioned in Section 4.1, the model places each development at its chosen prototype's lot size $x_{1,k}$, which may differ from the parcel's actual area A_i, raising the concern that the predicted supply gains rely on implausible land reconfiguration. Appendix Table A.1 reports the distribution of the ratio $x_{1,k}/A_i$ across expected development, where a ratio above one implies that development would require assembling adjacent land. Most development does not depend on assembly, as the median ratio is 0.82 at baseline and 87% of expected development on improved parcels occurs at or below the actual parcel area. The reforms that drive the largest supply gains shift development toward smaller lots and so rely less on assembly, with the median ratio falling to 0.60 and the improved-parcel share at or below the actual area rising to 96% under the comprehensive combined reform. The assembly that remains is concentrated on vacant parcels, which in Greater Boston are predominantly small infill remnants. This points to a limitation that grows more consequential going forward, since the remaining developable vacant land is scarce and fragmented into small parcels, so achieving an efficient lot increasingly requires assembly or lot-line adjustment that the model treats as costless.

5 Conclusion

This paper studies how land availability, construction costs, zoning regulations, and implicit barriers jointly shape single-family housing supply in Greater Boston. Using parcel-level data on zoning, transactions, assessments, and construction costs, I document substantial heterogeneity in local supply conditions across municipalities and show that observed development patterns are jointly determined by multiple interacting supply constraints rather than any single determinant.

I develop and estimate a discrete-choice model of single-family residential development in which landowners choose whether to develop and what type of home to build. The model combines observed prices, construction costs, existing structure values, and zoning constraints, and uses the gap between predicted profitability and observed development activity to recover municipality-specific implicit development costs. The estimates reveal substantial cross-municipality variation in these costs, which should be interpreted as capturing barriers beyond explicit permitting fees embedded in construction costs, including delay, uncertainty, and other soft costs of navigating the local approval process.

I then use the estimated model to examine how relaxing each constraint affects housing supply. The counterfactuals show that reducing implicit development costs, relaxing minimum lot size requirements, and lowering construction costs can all substantially increase housing production, with the largest effects arising when reforms are combined. Allowing prices to adjust in equilibrium attenuates these quantity effects but leaves the qualitative conclusions unchanged. The counterfactuals also imply important composition effects, as relaxing supply constraints shifts development toward smaller lots and denser housing types.

Overall, the results suggest that housing supply in high-demand metropolitan areas is constrained not by any single friction, but by the interaction of land scarcity, construction costs, formal zoning, and implicit barriers. At the same time, the analysis has several limitations. Importantly, it focuses on single-family development and does not account for multifamily or higher-density construction, so it captures only one component of metropolitan housing supply. The analysis also studies a single metropolitan area, abstracts from dynamic development timing, and treats zoning as time-invariant over the sample period, and the equilibrium analysis is deliberately parsimonious and does not allow for richer demand responses across housing types or neighborhoods. These limitations notwithstanding, the framework provides a tractable way to disentangle how multiple supply constraints jointly shape development outcomes and can be extended to multifamily development, richer demand models, and other geographic contexts.

Figures and Tables

Table 1 — Zoning regulations and bindingness

	N	Mean	SD	P10	P50	P90
Panel A. Dimensional requirements: single-family parcels (N = 468,088)						
Minimum lot size (sq ft)	447,380	26,664	30,162	7,500	15,500	45,000
Max floor-area ratio (FAR)	97,415	0.652	0.950	0.300	0.500	0.900
Panel B. Dimensional requirements: vacant parcels (N = 33,871)						
Minimum lot size (sq ft)	30,105	32,710	36,570	7,000	20,000	65,340
Max floor-area ratio (FAR)	9,055	0.858	1.377	0.300	0.600	1.500
Panel C. Bindingness among new construction (2010–2019)						
MLS binds ($A_i/\text{MLS}_i \in [1, 1.05]$)	5,319	0.129	–	–	–	–
MLS exists and binds	5,206	0.131	–	–	–	–
FAR binds ($\text{FAR}_i/\overline{\text{FAR}}_i \in [0.95, 1]$)	5,319	0.002	–	–	–	–
FAR exists and binds	1,256	0.010	–	–	–	–

Notes: Statistics computed on the analysis sample of 501,959 parcels across 98 municipalities. Panels A and B report dimensional zoning requirements assigned to parcels via spatial join between parcel centroids and zoning district polygons. FAR is merged from the MAPC Zoning Atlas; coverage is lower because many districts do not specify a maximum FAR. Panel C reports bindingness proxies among single-family homes built during 2010–2019, after trimming outliers using a 1.5× IQR rule on lot size and building area. A_i is lot area and MLS_i is the minimum lot size requirement. “Exists and binds” conditions on the regulation being non-missing.

Table 2 — Construction cost estimates by building type, location, and vintage

	2014	2019	2024	Pct. chg. 14–24	Pct. chg. 19–24
Panel A. Mean construction cost ($ per sq ft) by building-type bin					
Small (P0–P33 of sq ft)	150.3	150.9	176.3	0.58	16.6
Medium (P33–P67 of sq ft)	132.8	133.2	154.8	0.50	16.1
Large (P67–P100 of sq ft)	126.6	126.8	147.2	0.27	16.0
All prototypes	136.6	136.9	159.4	0.45	16.2
Panel B. Dispersion in construction costs ($ per sq ft)					
Q1	113.5	114.9	133.4	−1.56	13.00
Q2 (median)	134.7	133.5	153.3	−0.16	16.01
Q3	166.8	167.1	200.4	3.42	20.00
SD across municipalities	19.13	18.31	23.74	2.03	2.28
Mean within-municipality SD across prototypes	12.56	11.71	13.42	1.06	0.59

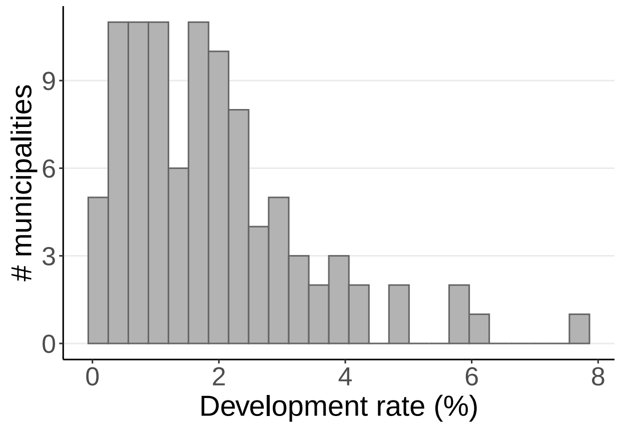
Notes: This table describes construction cost estimates from Craftsman Costbooks for single-family prototypes in the analysis sample. The unit of observation is a municipality-specific prototype home. Panel A reports mean costs in dollars per square foot by prototype size bin (terciles of pooled building square footage) and vintage. Panel B reports cross-municipality dispersion using quartiles and the standard deviation of municipality mean costs, as well as the average within-municipality standard deviation across prototypes.

Table 3 — Sample Statistics

	N	Mean	SD	P10	P50	P90
Panel A. Baseline parcel characteristics (as of 2009)						
Vacant parcel (indicator)	501,959	0.067	–	–	–	–
Lot size (sq ft)	501,941	16,795	13,264	4,356	12,000	40,075
Building area (sq ft, improved)	444,396	3,034	1,185	1,632	2,832	4,757
Year built (improved)	468,088	1,946	38	1,900	1,954	1,991
Panel B. Development outcomes (2010–2019)						
Development rate: all parcels	501,959	0.018	–	–	–	–
Development rate: vacant parcels	33,871	0.105	–	–	–	–
Development rate: improved parcels	468,088	0.012	–	–	–	–
Panel C. New construction transaction prices (2010–2019)						
Sale price ($)	19,431	712,461	562,469	200,000	570,000	1,432,000
Unit price ($ per bldg sq ft)	19,431	235	142	78	223	399
Panel D. Existing home transaction prices (2010–2019)						
Sale price ($)	143,990	451,074	196,696	230,000	415,000	740,000
Unit price ($ per bldg sq ft)	143,990	259	100	148	245	385
Panel E. Vacant land transaction prices (2010–2019)						
Sale price ($)	2,080	519,417	700,874	35,000	300,000	1,177,300
Unit price ($ per lot sq ft)	2,080	24.3	34.1	0.8	10.1	66.8

Notes: This table shows summary statistics for the analysis sample. Panel A reports baseline characteristics as of 2009. Panel B reports development rates over 2010–2019, where development is defined as new single-family construction identified from assessor records. Panels C–E report winsorized (1st and 99th percentile) arm's-length transaction prices from ATTOM deed records over 2010–2019, restricted to the 98 analysis sample municipalities.

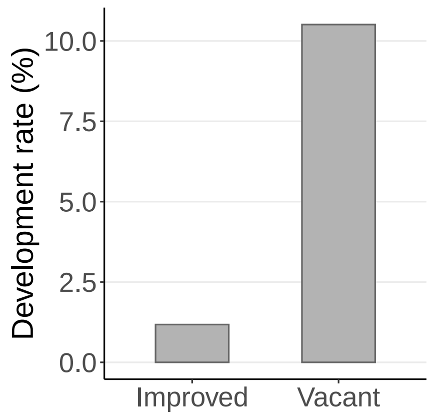
Figure 1 — Distribution of municipality-level development rates (2010–2019)



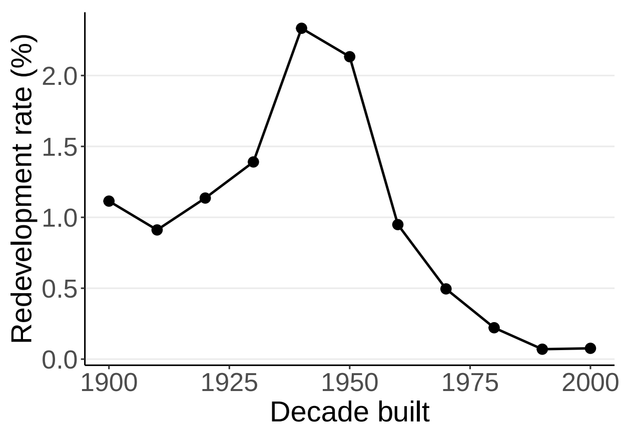
Note. This plot shows the histogram of municipality-level development rates. Each bar represents the number of municipalities with development rates in the corresponding bin. The development rate is the share of single-family and vacant parcels that are observed to develop a new single-family home over 2010–2019. The sample includes 98 municipalities in the Greater Boston metro.

Figure 2 — Development rates by existing parcel characteristics

(a) By parcel status



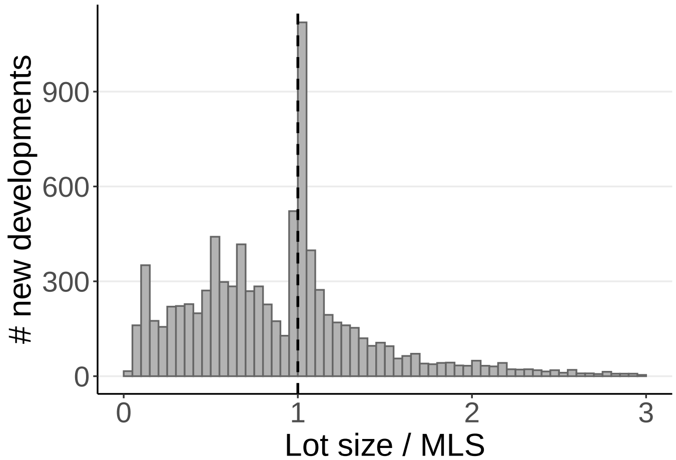
(b) By decade built (improved parcels)



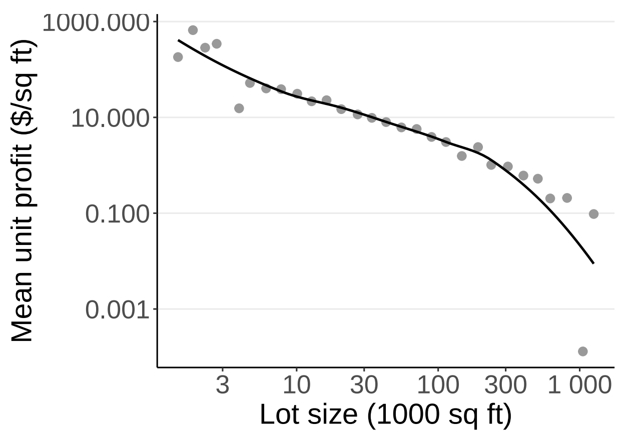
Note. Panel (a) plots the share of parcels that develop over 2010–2019, separately for parcels classified as vacant and improved as of 2009. Panel (b) plots the redevelopment rate among improved parcels by decade of original construction. Decades prior to 1900 are grouped together. The sample includes 468,088 improved parcels and 33,871 vacant parcels.

Figure 3 — Zoning constraints and development profitability by lot size

(a) Lot size relative to MLS



(b) Unit profit by lot size



Note. Panel (a) plots the distribution of lot size divided by the assigned minimum lot size (MLS) among parcels that developed during 2010–2019, restricted to parcels with non-missing MLS and lot-to-MLS ratios below 3. The dashed vertical line marks ratio = 1 (lot exactly meets MLS). Bin width is 0.05, aligned so that a bin edge falls at 1.0. Panel (b) plots mean unit profit (sale price minus construction cost, divided by lot size, in dollars per square foot) against lot size. Each point represents the mean within one of 30 equal-width bins in log lot size, and the lot-size axis is on a log scale. Bins with fewer than 100 observations or negative mean profit are excluded.

Table 4 — Municipality-level correlations among supply-side characteristics

	Dev. rate	Pct. vacant	Mean MLS	Mean price	Mean cost
Pct. vacant	0.07				
Mean MLS	0.30	0.14			
Mean price	0.24	−0.27	−0.03		
Mean cost	0.34	−0.11	0.16	0.73	
Mean profit	−0.19	−0.18	−0.31	0.50	0.06

Note. This table reports pairwise correlations across 98 municipalities. Dev. rate is the share of parcels developing over 2010–2019. Pct. vacant is the share of parcels classified as vacant in 2009. Mean MLS, price, cost, and profit are municipality-level averages across building types in the estimation sample. Profit is defined as sale price minus construction cost, divided by lot size (dollars per square foot of land).

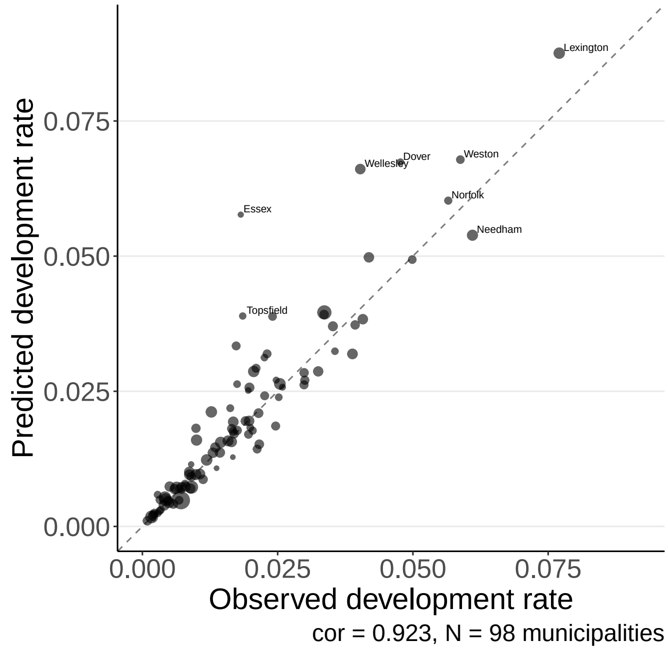
Table 5 — Summary of Estimated Implicit Development Cost Parameters

	N	Mean	Median	SD	p10	p90	Signif. (5%)
$\omega_{m,1}$ ($)	98	223,365	117,285	446,438	−53,341	510,382	89
$\omega_{m,2}$ ($)	95	58,781	44,959	268,056	−168,291	230,954	78

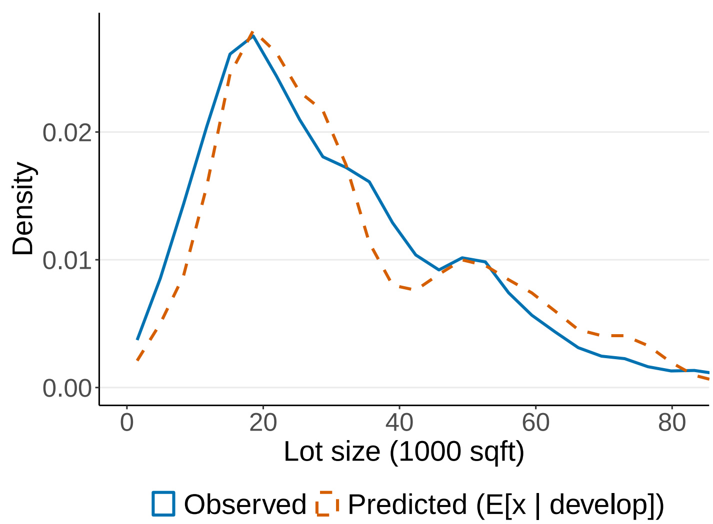
Note. This table describes the estimated implicit development cost parameters. Each municipality is estimated independently by maximum likelihood (BFGS with analytic gradients). $\omega_{m,1}$ is estimated for all 98 municipalities; $\omega_{m,2}$ is estimated for 95 municipalities where the choice set includes types both above and below the minimum lot size with non-degenerate development shares (3 municipalities have $\omega_{m,2}$ set to zero). Significance is based on the z-statistic from the observed Hessian.

Figure 4 — Model Fit

(a) Municipality development rates



(b) Lot-size distribution of developed parcels



Note. Panel (a) shows observed versus predicted municipality-level development rates across the 98 analysis-sample municipalities. Each point represents a municipality, point size is proportional to the number of parcels, and the dashed line is the 45-degree reference. Labeled municipalities are those with the largest absolute deviations from the 45-degree line or observed development rates above 6%. Panel (b) shows the density of observed lot sizes among developed parcels (solid) and model-predicted expected lot size conditional on development (dashed); the predicted distribution is computed from the conditional choice probabilities $Pr_i(k \mid \text{develop} = 1)$ weighted by the lot size of each building type.

Table 6 — Robustness of Estimated Implicit Development Costs Across Specifications

Specification	$\omega_{m,1}$ med. ($k)	$\omega_{m,2}$ med. ($k)	$\omega_{m,1}$ SD ($k)	ρ
Baseline (GAM-spatial, mean SQ, γ_0=$7)	117	45	446	0.923
Add explicit zoning controls	114	39	463	0.924
OLS prices & status quo	82	35	445	0.903
Period-minimum status quo	199	17	451	0.915
Low demolition cost (γ_0=$4/sqft)	121	45	446	0.923
High demolition cost (γ_0=$10/sqft)	116	40	447	0.922
Exclude imputed lot sizes	98	48	451	0.960
Min SQ + low demolition	201	15	451	0.915

Note. This table reports regulatory cost estimates under alternative specifications. Median and SD of $\omega_{m,1}$ and median of $\omega_{m,2}$ are computed across converged municipalities. ρ is the correlation between observed and predicted municipality-level development rates. *Baseline* uses GAM-spatial hedonic price indices for new construction and status quo values, the period-mean status quo value, and a demolition cost of $7 per square foot. *Add explicit zoning controls* augments the development cost with district-level floor-area ratio, maximum height, lot coverage, and a by-right multifamily indicator, with a single pooled coefficient on each estimated jointly with the municipality-specific costs (see Appendix D for details). *OLS prices and status quo* replaces the GAM-spatial hedonic models with OLS log-linear hedonic models for both new-construction prices and status quo values. *Period-minimum status quo* sets the status quo value to its minimum predicted level over 2010–2019 rather than the period mean. *Low* and *high demolition cost* set the per-square-foot demolition cost γ_0 to $4 and $10, respectively, instead of the $7 baseline. *Exclude imputed lot sizes* drops parcels whose lot size is imputed. *Min SQ + low demolition* combines the period-minimum status quo with the $4 demolition cost.

Table 7 — Counterfactual Housing Supply: Exogenous Prices

Scenario	Units	Δ%	% of stock	E[lot] (k sqft)
Baseline	14,466		2.88	40.1
High construction cost	7,657	−47.1	1.53	59.9
Low construction cost	25,896	+79.0	5.16	29.9
Remove MLS	18,419	+27.3	3.67	32.9
Reduce ω_1 \$50k (partial)	25,819	+78.5	5.14	27.8
Reduce ω_1 \$100k (comprehensive)	41,780	+188.8	8.32	21.9
Combined (Remove MLS + \$50k)	34,702	+139.9	6.91	23.4
Combined (Remove MLS + \$100k)	53,687	+271.1	10.70	19.6

Note. This table reports fixed-price counterfactual results. Units are predicted new single-family homes over the 2010–2019 sample period under the stated scenario. Δ% is the percent change in predicted units relative to the baseline scenario. % of stock expresses predicted units as a share of the 501,959-parcel analysis sample over the period. $E[\text{lot}]$ is the expected lot size among new developments (thousand sqft). The high construction cost scenario uses 2024 Craftsman estimates and the low construction cost scenario applies a 15% discount to baseline costs. MLS denotes minimum lot size. The partial reform reduces $\omega_{m,1}$ by \$50,000 and the comprehensive reform reduces it by \$100,000. Combined reforms remove MLS and reduce $\omega_{m,1}$ by the stated amount.

Table 8 — Counterfactual Housing Supply: Equilibrium Prices

Panel A ($\eta = 0.3$)					
Scenario	Units	Δ%	% of stock	E[lot] (k sqft)	$\bar{P}_m$
High construction cost	8,273	−42.8	1.65	55.8	1.02
Low construction cost	24,714	+70.8	4.92	30.8	0.99
Remove MLS	18,607	+28.6	3.71	32.1	1.00
Reduce ω_1 \$50k (partial)	25,002	+72.8	4.98	28.3	0.99
Reduce ω_1 \$100k (comprehensive)	39,395	+172.3	7.85	22.5	0.98
Combined (Remove MLS + \$50k)	33,190	+129.4	6.61	23.6	0.98
Combined (Remove MLS + \$100k)	50,988	+252.5	10.16	19.8	0.95

Panel B ($\eta = 0.1$)					
Scenario	Units	Δ%	% of stock	E[lot] (k sqft)	$\bar{P}_m$
High construction cost	7,928	−45.2	1.58	58.0	1.01
Low construction cost	25,475	+76.1	5.08	30.2	1.00
Remove MLS	18,405	+27.2	3.67	32.8	1.00
Reduce ω_1 \$50k (partial)	25,551	+76.6	5.09	27.9	1.00
Reduce ω_1 \$100k (comprehensive)	40,929	+182.9	8.15	22.1	0.99
Combined (Remove MLS + \$50k)	34,069	+135.5	6.79	23.5	0.99
Combined (Remove MLS + \$100k)	52,681	+264.2	10.50	19.7	0.98

Panel C ($\eta = 0.6$)					
Scenario	Units	Δ%	% of stock	E[lot] (k sqft)	$\bar{P}_m$
High construction cost	8,616	−40.4	1.72	53.9	1.03
Low construction cost	23,873	+65.0	4.76	31.4	0.98
Remove MLS	18,774	+29.8	3.74	31.5	1.00
Reduce ω_1 \$50k (partial)	24,254	+67.7	4.83	28.9	0.98
Reduce ω_1 \$100k (comprehensive)	37,508	+159.3	7.47	23.1	0.96
Combined (Remove MLS + \$50k)	32,038	+121.5	6.38	23.8	0.96
Combined (Remove MLS + \$100k)	48,490	+235.2	9.66	20.0	0.92

Note. This table reports counterfactual results with endogenous price responses, after iterating to a fixed point of Equation (9). Units are predicted new single-family homes over the 2010–2019 sample period under the stated scenario. Baseline units are 14,466. Δ% is the percent change in predicted units relative to the baseline scenario. % of stock and $E[\text{lot}]$ are defined as in Table 7. $\bar{P}_m$ is the mean equilibrium price multiplier across municipalities, where 1.00 denotes observed prices. Scenarios match Table 7. MLS denotes minimum lot size.

References

Ahlfeldt, Gabriel and Daniel P McMillen, "New estimates of the elasticity of substitution of land for capital," 2014.

Anagol, Santosh, Fernando Ferreira, and Jonah Rexer, "Estimating the Economic Value of Zoning Reform," *Working Paper*, 2021.

Baum-Snow, Nathaniel and Gilles Duranton, "Housing supply and housing affordability," in "Handbook of Regional and Urban Economics," Vol. 6, Elsevier, 2025, pp. 353–461.

_ and Lu Han, "The microgeography of housing supply," *Journal of Political Economy*, 2024, *132* (6), 1897–1946.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, "The production function for housing: Evidence from France," *Journal of Political Economy*, 2021, *129* (10), 2766–2816.

Davis, Morris A, William D Larson, Stephen D Oliner, and Jessica Shui, "The price of residential land for counties, ZIP codes, and census tracts in the United States," *Journal of Monetary Economics*, 2021, *118*, 413–431.

Dobbels, Gregory and Suren Tavakalov, "Not in My Back Yard: The Local Political Economy of Residential Land-Use Regulations," 2024.

Fang, Limin, Nathan Stewart, and Justin Tyndall, "Homeowner politics and housing supply," *Journal of Urban Economics*, 2023, *138*, 103608.

Favilukis, Jack Y and Jaehee Song, "Why zoning is too restrictive," *Available at SSRN 4535417*, 2023.

Gabriel, Stuart and Edward Kung, "Development Approval Times and New Housing Supply: Evidence from Los Angeles," *Available at SSRN 4872147*, 2025.

Glaeser, Edward and Joseph Gyourko, "The economic implications of housing supply," *Journal of economic perspectives*, 2018, *32* (1), 3–30.

Glaeser, Edward L and Bryce A Ward, "The Causes and Consequences of Land Use Regulation: Evidence from Greater Boston," *Journal of urban Economics*, 2009, *65* (3), 265–278.

_ and Joseph Gyourko, "America's Housing Supply Problem: The Closing of the Suburban Frontier?," Technical Report, National Bureau of Economic Research 2025.

_ , _ , and Raven E Saks, "Why have housing prices gone up?," *American Economic Review*, 2005, *95* (2), 329–333.

Green, Richard K., Stephen Malpezzi, and Stephen K. Mayo, "Metropolitan-Specific Estimates of the Price Elasticity of Supply of Housing, and Their Sources," *American Economic Review*, 2005, *95* (2), 334–339.

Gyourko, Joseph, Albert Saiz, and Anita Summers, "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index," *Urban Studies*, 2008, *45* (3), 693–729.

— **and Jacob Krimmel**, "The Impact of Local Residential Land Use Restrictions on Land Values across and within Single Family Housing Markets," *Journal of Urban Economics*, 2021, *126*, 103374.

— **and Raven Molloy**, "Regulation and housing supply," in "Handbook of regional and urban economics," Vol. 5, Elsevier, 2015, pp. 1289–1337.

He, Dongyang, "Distributional Effects of Residential Zoning Policies: Insights from the Greater Boston Area," *Available at SSRN 6172940*, 2024.

Hilber, Christian AL and Frédéric Robert-Nicoud, "On the origins of land use regulations: Theory and evidence from US metro areas," *Journal of Urban Economics*, 2013, *75*, 29–43.

Howard, Troup, Mengqi Wang, and Dayin Zhang, "Cracking Down, Pricing Up: Housing Supply in the Wake of Mass Deportation," *Pricing Up: Housing Supply in the Wake of Mass Deportation*, 2024.

Jackson, Kristoffer, "Do land use regulations stifle residential development? Evidence from California cities," *Journal of urban economics*, 2016, *91*, 45–56.

Kulka, Amrita, Aradhya Sood, and Nicholas Chiumenti, "Under the (neighbor) hood: Understanding interactions among zoning regulations," *Review of Economics and Statistics*, 2026, pp. 1–45.

Li, Xiaodi, "Do new housing units in your backyard raise your rents?," *Journal of Economic Geography*, 2022, *22* (6), 1309–1352.

Ma, Lei, "Build What and for Whom? The Distributional Effects of Housing Supply," *The Distributional Effects of Housing Supply (October 01, 2024)*, 2024.

Mast, Evan, "Warding off development: Local control, housing supply, and nimbys," *Review of Economics and Statistics*, 2024, *106* (3), 671–680.

Mayer, Christopher J and C Tsuriel Somerville, "Land use regulation and new construction," *Regional Science and Urban Economics*, 2000, *30* (6), 639–662.

Mense, Andreas, "The impact of new housing supply on the distribution of rents," *Journal of Political Economy Macroeconomics*, 2025, *3* (1), 1–42.

Molloy, Raven, Charles G Nathanson, and Andrew Paciorek, "Housing supply and affordability: Evidence from rents, housing consumption and household location," *Journal of Urban Economics*, 2022, *129*, 103427.

Murphy, Alvin, "A dynamic model of housing supply," *American economic journal: economic policy*, 2018, *10* (4), 243–267.

Paciorek, Andrew, “Supply Constraints and Housing Market Dynamics,” *Journal of Urban Economics*, 2013, *77*, 11–26.

Rollet, Vincent, “Zoning and the Dynamics of Urban Redevelopment,” 2025.

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

Soltas, Evan and Jonathan Gruber, “How Costly Is Permitting in Housing Development?,” 2026.

Song, Jaehee, “The effects of residential zoning in US housing markets,” *Journal of Urban Economics*, 2025, *149*, 103784.

Wrenn, Douglas H and Elena G Irwin, “Time is money: An empirical examination of the effects of regulatory delay on residential subdivision development,” *Regional Science and Urban Economics*, 2015, *51*, 25–36.

Appendix

A Appendix Tables and Figures

Figure A.1 — Example Craftsman Costbooks cost report (direct cost items)

Replacement Estimate by Cost Category

Direct Cost Items				
Item Name	**Materials**	**Labor**	**Equipment**	**Total**
Excavation	XXX	XXX	XXX	XXX
Foundation, Piers, Flatwork	XXX	XXX	XXX	XXX
Insulation	XXX	XXX	XXX	XXX
Rough Hardware	XXX	XXX	XXX	XXX
Framing	XXX	XXX	XXX	XXX
Exterior Finish	XXX	XXX	XXX	XXX
Exterior Trim	XXX	XXX	XXX	XXX
Doors	XXX	XXX	XXX	XXX
Windows	XXX	XXX	XXX	XXX
Roofing, Soffit, Fascia	XXX	XXX	XXX	XXX
Finish Carpentry	XXX	XXX	XXX	XXX
Interior Wall Finish	XXX	XXX	XXX	XXX
Lighting Fixtures	XXX	XXX	XXX	XXX
Painting	XXX	XXX	XXX	XXX
Carpet, Flooring	XXX	XXX	XXX	XXX
Bath Accessories	XXX	XXX	XXX	XXX
Shower & Tub Enclosures	XXX	XXX	XXX	XXX
Plumbing Fixtures	XXX	XXX	XXX	XXX
Plumbing Rough-in	XXX	XXX	XXX	XXX
Wiring	XXX	XXX	XXX	XXX
Built In Appliances	XXX	XXX	XXX	XXX
Cabinets	XXX	XXX	XXX	XXX
Countertops	XXX	XXX	XXX	XXX
Central Heating and Cooling	XXX	XXX	XXX	XXX
Fireplace	XXX	XXX	XXX	XXX
Subtotal direct job cost	XXX	XXX	XXX	XXX

Note. This figure shows an example Craftsman Costbooks cost report listing direct cost items. Dollar amounts are masked due to data licensing restrictions.

Figure A.2 — Example Craftsman Costbooks cost report (indirect cost items and total)

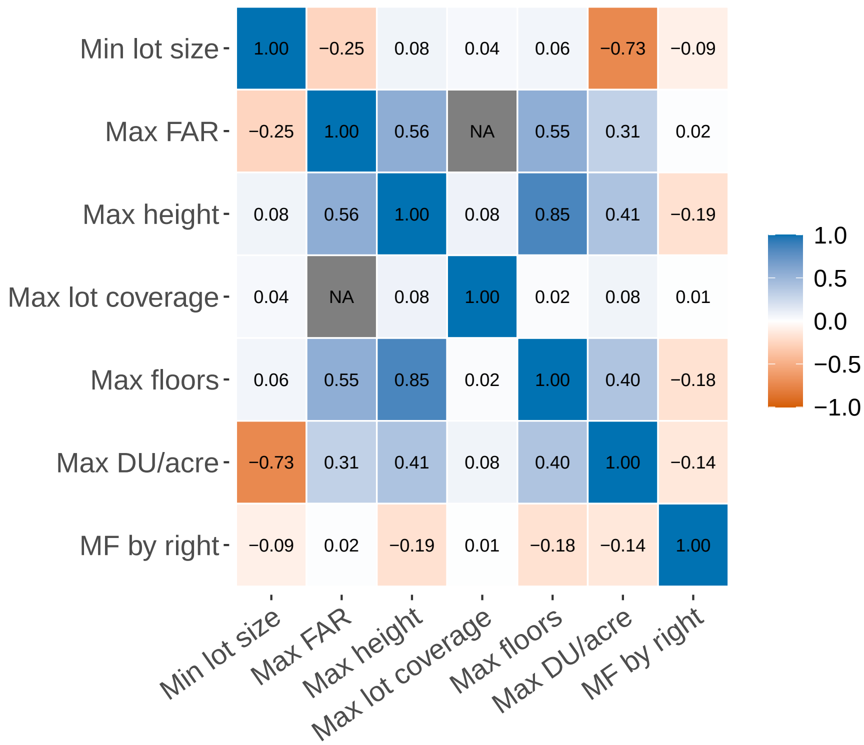
Indirect Cost Items				
Item Name	**Materials**	**Labor**	**Equipment**	**Total**
Final Cleanup	XXX	XXX	XXX	XXX
Insurance	XXX	XXX	XXX	XXX
Permits & Utilities	XXX	XXX	XXX	XXX
Design & Engineering	XXX	XXX	XXX	XXX
Subtotal indirect job cost	XXX	XXX	XXX	XXX

Grand Total				
Item Name	**Materials**	**Labor**	**Equipment**	**Total**
Contractor Markup	XXX	XXX	XXX	XXX
Total cost	XXX	XXX	XXX	XXX

This replacement cost estimate is based on the same sources and methods used to compile figures which appear in *2014 National Building Cost Manual* published annually by Craftsman Book Company, 6058 Corte del Cedro, Carlsbad, CA 92011, 1-800-829-8123, http://www.craftsman-book.com. The labor, material and equipment costs in this report have been adjusted by factors in a cost index for the Zip area XXX XXX Current index factors for this area are XXX for material, XXX or labor anc XXX for equipment. This estimate assumes a single home is being replaced.

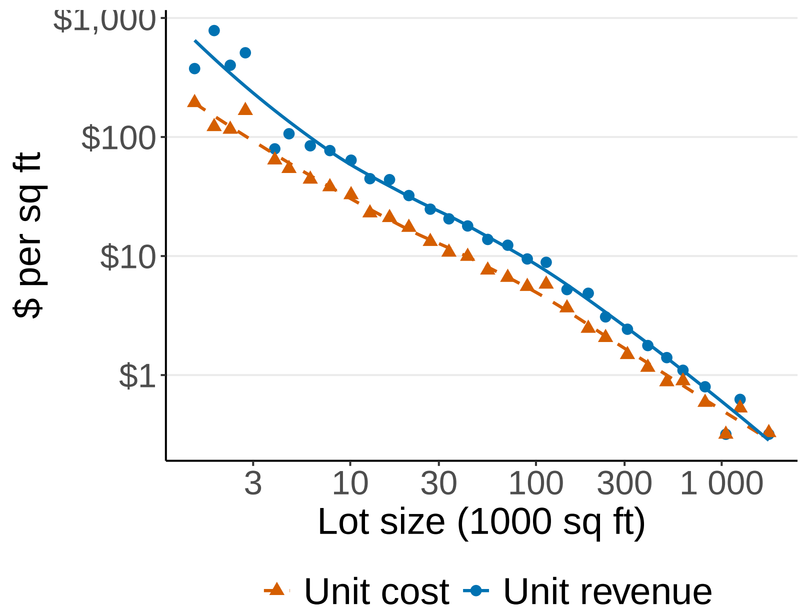
Note. This figure shows an example Craftsman Costbooks cost report listing indirect cost items and total cost estimate. Dollar amounts and factor multipliers are masked due to data licensing restrictions.

Figure A.3 — Co-variation among zoning regulations



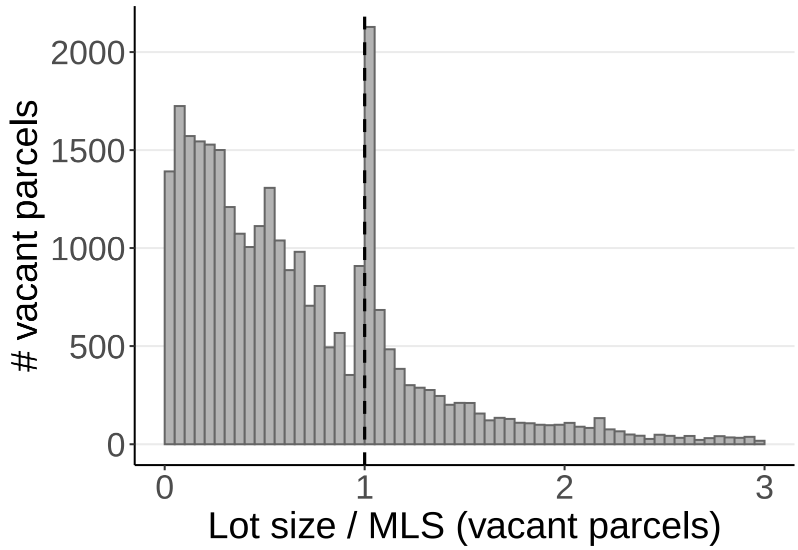
Note. The figure depicts Spearman rank correlations across zoning districts among minimum lot size, maximum floor-area ratio, maximum height, maximum lot coverage, maximum floors, maximum dwelling units per acre, and an indicator for whether multifamily housing is allowed by right. Effective (harmonized) regulatory values are used, and districts where a regulation is absent are treated as missing, so correlations use pairwise-complete observations.

Figure A.4 — Revenue and cost per square foot of land by lot size



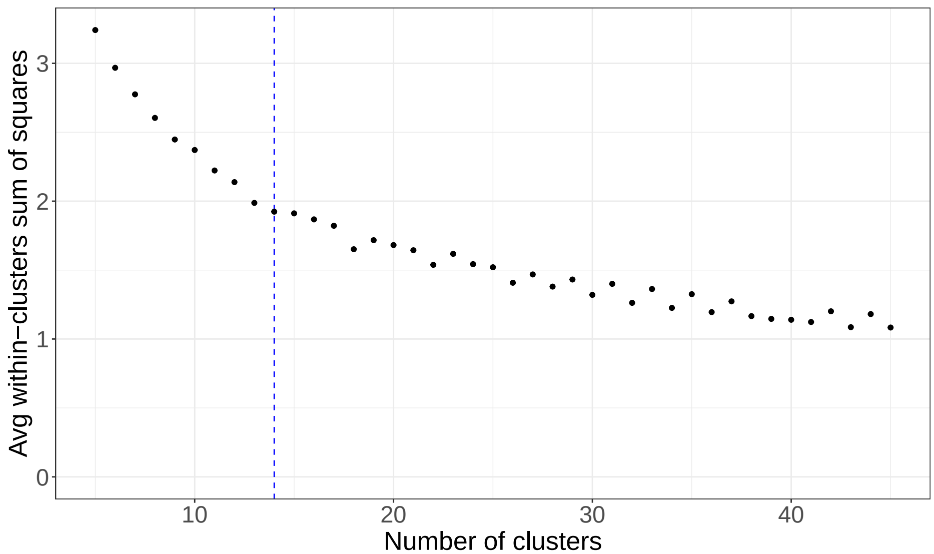
Note. This plot shows mean unit revenue (sale price per square foot, blue circles) and mean unit cost (construction cost per square foot, orange triangles) against lot size. Each point represents the mean within one of 30 equal-width bins in log lot size, and both axes are on a log scale. Bins with fewer than 100 observations are excluded.

Figure A.5 — Lot size relative to minimum lot size among vacant parcels



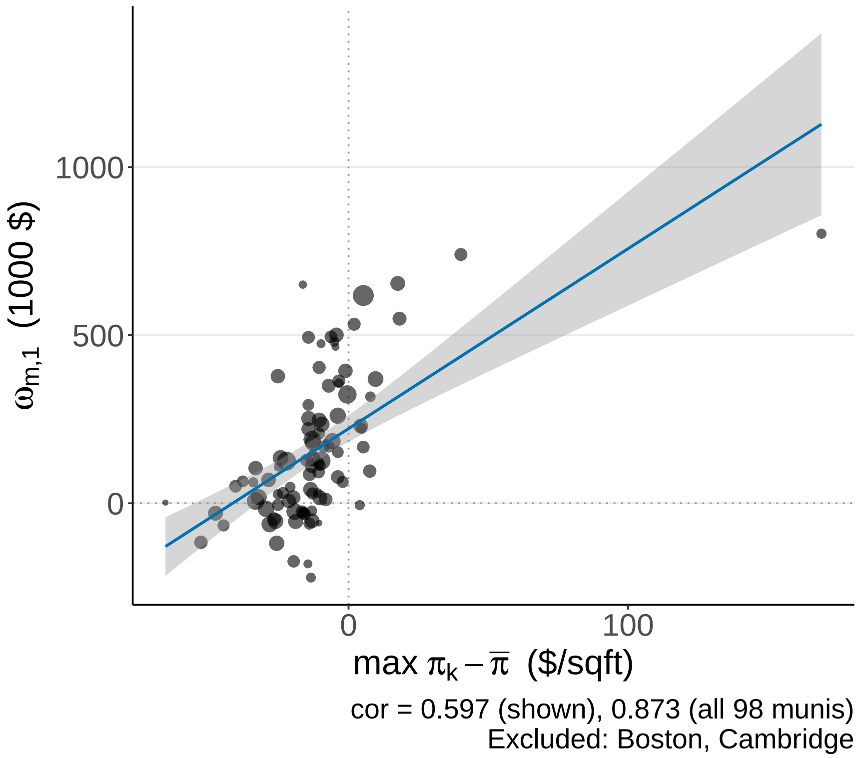
Note. This figure plots the distribution of lot size divided by the assigned minimum lot size (MLS) among vacant parcels, restricted to parcels with non-missing MLS and ratios below three. The dashed vertical line marks a ratio of one.

Figure A.6 — Within-Cluster Sum of Squared Errors by Number of Clusters Using K-Prototypes



Note. The figure plots the within-cluster sum of squared errors (WSS) for different values of k in the k-prototypes clustering algorithm as described in Section 4.2. The selected number of clusters, $k = 14$, is depicted by a blue dashed line.

Figure A.7 — Estimated Implicit Development Cost vs. Development Profitability



Note. Each point represents a municipality. The horizontal axis is the gap between the maximum unit profit across building types and the mean status quo value ($/sqft); positive values indicate that development is profitable before regulatory costs. The vertical axis is the estimated $\omega_{m,1}$. Point size is proportional to the number of parcels. The line is an OLS fit with a 95% confidence band. Boston ($3.4M) and Cambridge ($1.7M) are excluded from the plot to preserve scale; including them raises the correlation from 0.597 to 0.87.

Table A.1 — Counterfactual chosen lot size relative to actual parcel size

		Share of expected development by $x_{1,k}/A_i$ (%)			
Scenario	Parcels	≤ 1	$(1, 1.5]$	$(1.5, 2]$	> 2
Baseline	All	61.1	13.2	6.0	19.8
	Improved	87.0	11.0	1.5	0.5
	Vacant	25.2	16.2	12.2	46.4
High construction cost	All	61.5	11.7	5.4	21.3
	Improved	88.4	10.0	1.2	0.4
	Vacant	23.7	14.2	11.4	50.7
Remove MLS	All	67.0	10.0	6.1	16.9
	Improved	94.0	4.4	1.1	0.4
	Vacant	37.5	16.1	11.5	34.9
Reduce $\omega_{m,1}$ \$100k	All	68.4	12.0	5.7	14.0
	Improved	90.7	8.0	1.1	0.2
	Vacant	35.0	17.9	12.5	34.5
Combined (Remove MLS + \$100k)	All	76.9	7.8	4.3	11.0
	Improved	95.5	3.7	0.6	0.2
	Vacant	40.1	15.9	11.6	32.4

Note. Each cell reports the share of expected new development over 2010–2019 falling in a bin of the ratio of chosen prototype lot size $x_{1,k}$ to the parcel's actual area A_i. Expected development weights each parcel-by-type pair by $Pr_i(\text{develop}) \times Pr_i(k \mid \text{develop})$. A ratio at or below one indicates development uses no more than the actual parcel; a ratio above one implies land assembly. Parcels with missing actual area are excluded. Scenarios correspond to the fixed-price counterfactuals in Table 7. The expected-development median ratio is 0.82 at baseline and 0.60 under the combined reform.

B Hedonic Price Models

This appendix reports model fit diagnostics for the two hedonic price models used to construct inputs to the structural model: the new construction price index P_{ik} (Section B.2) and the status quo value $\bar{\pi}_i$ for improved parcels (Section B.3). Both models share the same core specification and are estimated across the same four variants; the key difference is that the status quo model adds building age as a predictor. I first describe the common structure, then report specification-specific results for each model.

B.1 Specification Overview

Both models are estimated on log transaction prices of single-family homes in the MAPC region over 2010–2019, winsorized at the 1st and 99th percentiles. Housing characteristics enter identically in both models: penalized splines on building square footage, lot size, finished basement area, and unfinished basement area (each paired with a binary indicator for structural zeros, since zero basement area indicates the absence of a basement rather than a small one), with the number of bathrooms, bedrooms, and stories entering as factors. Year×municipality fixed effects and calendar month of transaction are included in all specifications.

I estimate four variants that differ along two dimensions:

Functional form. OLS log-linear (continuous predictors enter linearly) versus GAM with penalized splines (continuous predictors enter as smooth functions estimated by fREML). The GAM permits non-linear marginal effects, for example, a concave relationship between building area and log price, while the OLS imposes linearity in all continuous covariates.

Location control. Zoning-district fixed effects (one intercept per district, absorbed via the within transformation in OLS) versus a continuous spatial surface over zoning-district centroid coordinates. In OLS, the spatial surface is a second-order polynomial in longitude and latitude (lon, lat, lon^2, lat^2, lon $\times$ lat); in GAM, it is a two-dimensional thin-plate regression spline $g(\text{lon}, \text{lat})$. The fixed-effect specifications absorb fine-grained location heterogeneity but cannot predict prices for zoning districts with no observed transactions. The spatial specifications trade some in-sample precision for the ability to extrapolate to unobserved districts.

The resulting four models are denoted OLS-FE, GAM-FE, OLS-spatial, and GAM-spatial. The baseline specification used in the structural model is the GAM-spatial variant. The GAM-FE model is estimated in two steps to avoid collinearity between the large number of zoning-district indicators and the penalized smooth terms: I first absorb zoning-district fixed effects via OLS (using `feols`), then estimate the GAM smooth terms on the residualized log price. The combined prediction sums the two components. Because this sequential procedure does not jointly optimize the smooth terms and the fixed effects, the GAM-FE's in-sample R^2 may be lower than the OLS-FE despite achieving better fit on other metrics such as MAE. For all four models, predicted prices are retransformed from the log scale using the standard smearing correction $\hat{P} = \exp(\hat{y} + \frac{1}{2}\hat{\sigma}^2)$, where $\hat{\sigma}^2$ is the estimated residual variance.

B.2 New Construction Price Index

Sample

The analysis sample consists of 19,431 transactions of newly built single-family homes across 98 municipalities and 483 zoning districts, with a median of 126 transactions per municipality and 12 per zoning district. Winsorized transaction prices range from $65,300 to $3,198,500.

In-sample fit comparison

Table A.2 reports R^2, root mean squared error (RMSE), and mean absolute error (MAE) on log-price residuals for all four models.

Table A.2 — In-Sample Fit: New Construction Price Models

Model	Location control	N	R^2	RMSE	MAE
OLS-FE	Zoning district FE	19,431	0.570	0.495	0.400
GAM-FE	Zoning district FE	19,431	0.544	0.510	0.395
OLS-spatial	Centroid surface	19,431	0.504	0.532	0.443
GAM-spatial	Centroid surface	19,431	0.537	0.514	0.418

Note. All models are estimated on log transaction prices of newly built single-family homes sold within two years of construction, 2010–2019. RMSE and MAE are computed on log-price residuals.

The FE models achieve higher in-sample R^2 than the spatial models, as zoning-district indicators absorb fine-grained location heterogeneity that a smooth surface cannot fully replicate. Within the spatial specifications, the GAM substantially improves on OLS (R^2 of 0.537 vs. 0.504; RMSE of 0.514 vs. 0.532), confirming that the non-linear smooth terms capture meaningful variation in the price–characteristics relationship. The GAM-FE achieves the lowest MAE (0.395) among all four models despite a lower R^2 than OLS-FE, consistent with the sequential estimation procedure described in Section B.1. However, the FE specifications cannot predict prices for zoning districts with no observed new-construction transactions, which motivates the spatial baseline for the structural model.

GAM smooth term diagnostics

Figure A.8 plots the estimated smooth terms from the GAM-spatial model. The purpose of these plots is to verify that the fitted smooth functions are well-behaved and to assess the degree of non-linearity captured by the GAM relative to OLS; because the model is used to construct price indexes rather than to identify causal effects, I do not interpret the partial effects structurally. Building square footage exhibits a strongly non-linear, concave relationship with log price (edf = 5.96), rising steeply up to approximately 4,000 square feet and flattening thereafter. Lot size is penalized to a near-linear effect (edf = 1.00) with a weakly negative slope. Finished basement area shows a mildly non-linear effect (edf = 3.65) that is negative on net, while unfinished basement area contributes a small effect with mild non-linearity (edf = 1.76). The two-dimensional spatial surface $g(\text{lon}, \text{lat})$ uses 25.35 effective degrees of freedom, capturing substantial location premia across the Greater Boston area. Table A.3 reports the full set of smooth term diagnostics.

Table A.3 — GAM-Spatial Smooth Term Diagnostics: New Construction

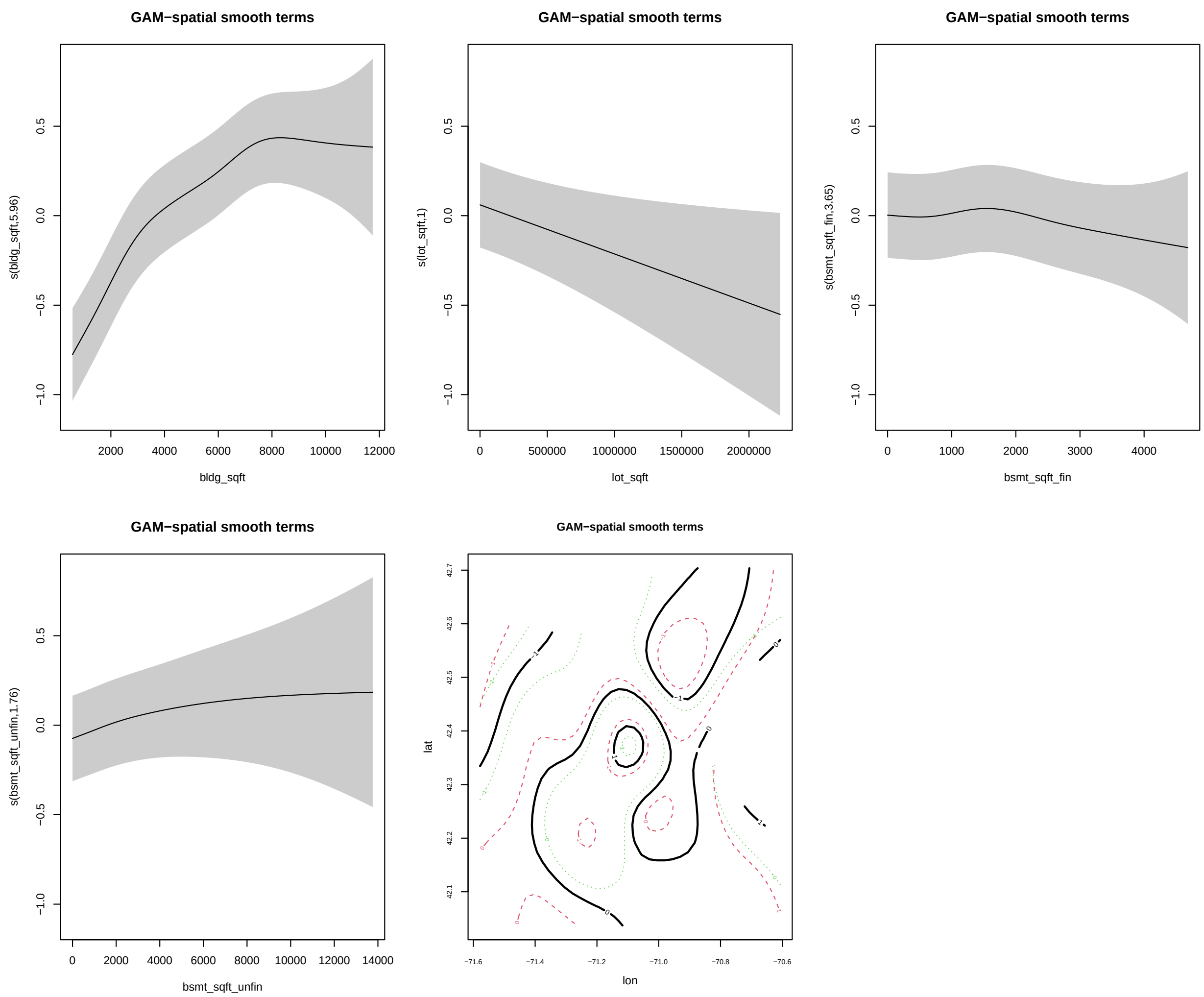
Smooth term	EDF	Ref. df	F	p-value
s(building sqft)	5.96	7.07	104.62	< 0.001
s(lot sqft)	1.00	1.00	5.36	0.021
s(fin. basement)	3.65	4.60	2.64	0.035
s(unfin. basement)	1.76	2.18	5.28	0.004
s(lon, lat)	25.35	28.08	20.67	< 0.001

Note. EDF is effective degrees of freedom (higher values indicate greater non-linearity; EDF = 1 corresponds to a linear effect). Ref. df is the reference degrees of freedom used for the approximate F-test. All terms are from the GAM-spatial specification estimated by fREML.

Observed vs. predicted prices

Figure A.9 compares observed and predicted transaction prices, aggregated at the municipality×zoning district×building type level, for all four models on a log scale. All four specifications cluster tightly around the 45-degree line across the full price range (\$60k–\$3.3M), with no systematic over- or under-prediction at either tail. The GAM-spatial and GAM-FE panels show slightly less dispersion than their OLS counterparts, consistent with the RMSE improvements reported in Table A.2.

Figure A.8 — GAM-Spatial Smooth Terms: New Construction



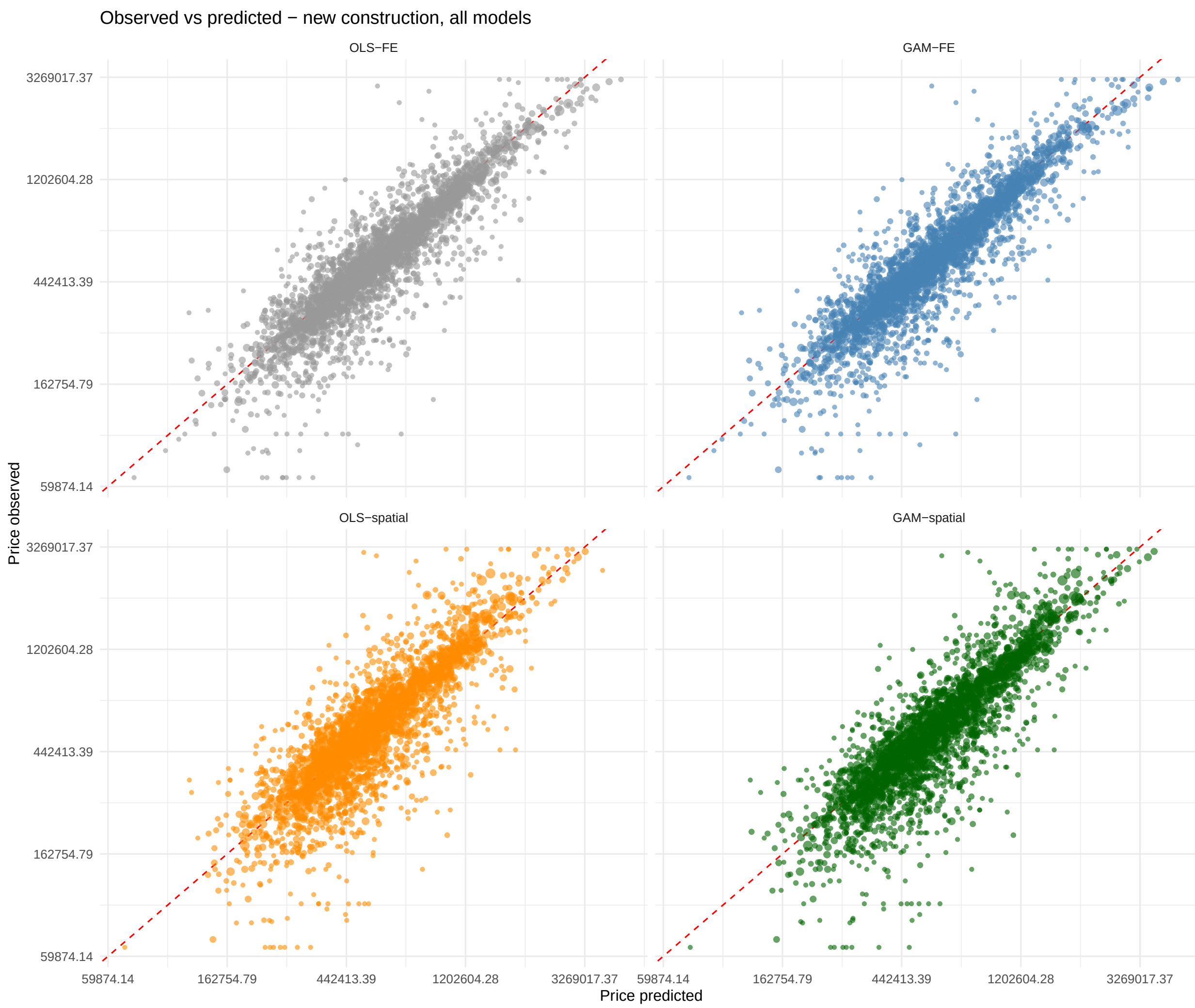
Note. Estimated smooth terms from the GAM-spatial model for new construction prices. Shaded regions represent approximate 95% confidence bands. The bottom panel shows the spatial surface $g(\text{lon}, \text{lat})$ over zoning-district centroids as a contour plot, with values in log-price units relative to the intercept.

B.3 Status Quo Value for Improved Parcels

Sample

The analysis sample consists of 143,328 transactions of existing single-family homes (at least two years old at time of sale) across 98 municipalities and 800 zoning districts, 2010–2019. This sample is roughly seven times larger than the new construction sample, with a median of 1,204 transactions per municipality and 15 per zoning district. Winsorized transaction prices range from \$96,000 to \$990,000. Building (effective) ages range from 2 to 319 years, with a median of 55.

Figure A.9 — Observed vs. Predicted Prices: New Construction



Note. Each point represents the mean observed vs. mean predicted transaction price for a municipality×zoning district×building type cell. Both axes on log scale. The dashed red line is the 45-degree reference. Panels show the four hedonic specifications.

In-sample fit comparison

Table A.4 reports R^2, RMSE, and MAE for all four status quo models.

All four models achieve substantially higher R^2 than the new construction models (0.65–0.68 vs. 0.50–0.57), reflecting the larger sample and the additional explanatory power of building age. The GAM-spatial specification achieves the highest R^2 (0.683), lowest RMSE (0.260), and lowest MAE (0.177) among all four models, confirming the value of both the non-linear smooth terms and the flexible depreciation curve relative to OLS.

Table A.4 — In-Sample Fit: Status Quo Price Models

Model	Location control	N	R^2	RMSE	MAE
OLS-FE	Zoning district FE	143,328	0.670	0.276	0.188
GAM-FE	Zoning district FE	143,328	0.668	0.277	0.182
OLS-spatial	Centroid surface	143,328	0.650	0.274	0.202
GAM-spatial	Centroid surface	143,328	0.683	0.260	0.177

Note. All models are estimated on log transaction prices of existing single-family homes (at least two years old at sale), 2010–2019. The status quo models include building age as an additional predictor (linear in OLS, penalized spline in GAM).

GAM smooth term diagnostics

The key difference from the new construction model is the addition of building age as a predictor. In the OLS specifications, effective age (computed as the transaction year minus the effective year built, which accounts for major renovations) enters linearly; in the GAM specifications, it enters as a penalized spline $s(\text{age}_{\text{eff}})$, allowing for a flexible depreciation curve. Figure A.10 plots the estimated smooth terms from the GAM-spatial status quo model. The age smooth (edf = 8.07) shows rapid depreciation over the first 50 years followed by a gradual flattening beyond age 100, consistent with the well-documented non-linear depreciation profile of residential structures. Building square footage shows a strongly concave effect (edf = 8.75), similar in shape to the new construction model. Lot size, in contrast to the new construction model where it was penalized to a linear effect, now exhibits substantial non-linearity (edf = 8.62), likely reflecting the much larger sample and greater variation in lot sizes among existing homes. Finished and unfinished basement area show moderately non-linear effects (edf = 5.43 and 5.80, respectively). The spatial surface captures location premia across the Greater Boston area with a pattern similar to the new construction model.

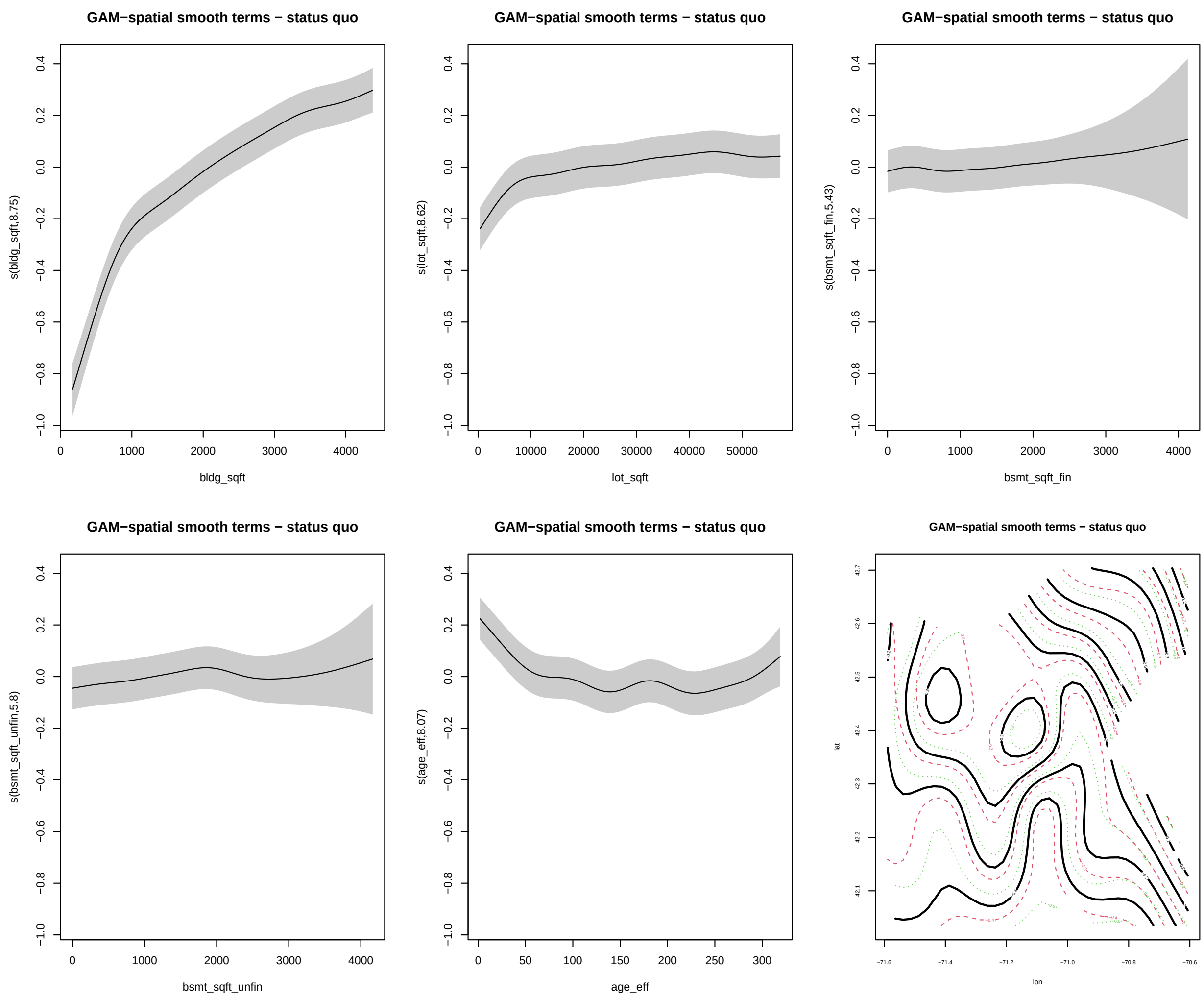
Observed vs. predicted prices

Figure A.11 compares observed and predicted transaction prices, aggregated at the municipality×zoning district level, for all four status quo models on a log scale. All four specifications cluster tightly around the 45-degree line across the price range ($120k–$885k), with noticeably less dispersion than the new construction models, reflecting the larger sample and greater predictive power of the age-augmented specification.

Out-of-sample prediction on baseline parcels

The fitted status quo models are applied to the inventory of pre-2009 improved parcels to predict each parcel's market value at each year from 2010 through 2019. For the structural model,

Figure A.10 — GAM-Spatial Smooth Terms: Status Quo



Note. Estimated smooth terms from the GAM-spatial model for existing home prices. Shaded regions represent approximate 95% confidence bands. The age panel shows the estimated depreciation curve. The bottom panel shows the spatial surface $g(\text{lon}, \text{lat})$ as a contour plot.

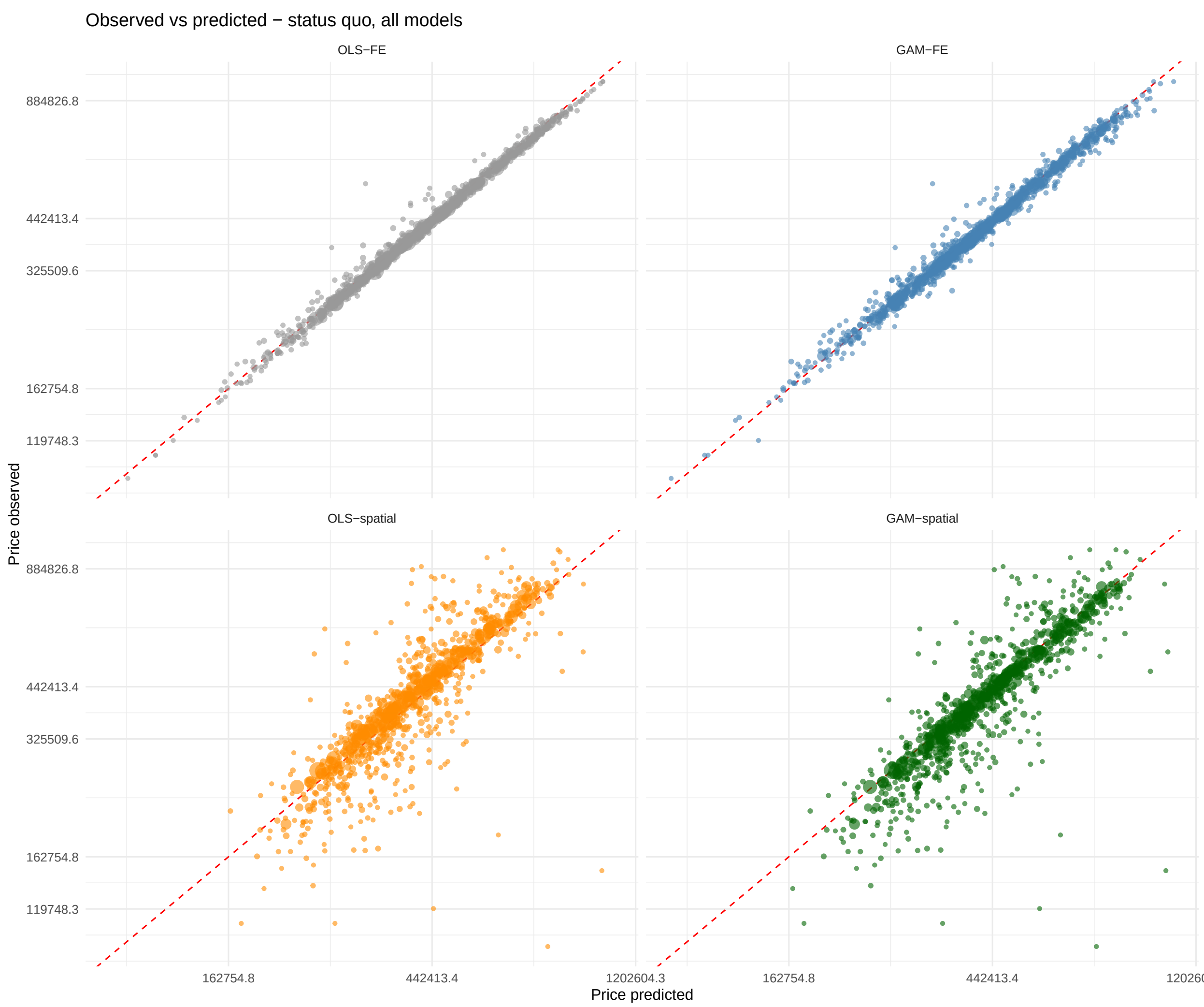
these annual predictions are averaged to obtain a single period-mean status quo value per parcel (the period minimum is used as a robustness check, as described in Section 4.1). The hedonic-predicted total home value is converted to a per-square-foot-of-land measure by dividing by the parcel's lot area. Missing lot area and building area are imputed using zoning-district-level medians from the non-missing parcels in the same district. Approximately 12% of parcels require imputation.

Table A.5 — GAM-Spatial Smooth Term Diagnostics: Status Quo

Smooth term	EDF	Ref. df	F	p-value
s(building sqft)	8.75	8.98	1,183.30	< 0.001
s(lot sqft)	8.62	8.95	251.85	< 0.001
s(fin. basement)	5.43	6.39	5.40	< 0.001
s(unfin. basement)	5.80	6.73	26.52	< 0.001
$s(\text{age}_{\text{eff}})$	8.07	8.62	633.86	< 0.001
s(lon, lat)	27.74	28.88	103.00	< 0.001

Note. EDF is effective degrees of freedom. All terms are from the GAM-spatial status quo specification estimated by fREML. Compared to the new construction model (Table A.3), all smooth terms use substantially more degrees of freedom, reflecting greater heterogeneity in the existing housing stock.

Figure A.11 — Observed vs. Predicted Prices: Status Quo



Note. Each point represents the mean observed vs. mean predicted transaction price for a municipality×zoning district cell. Both axes on log scale. The dashed red line is the 45-degree reference. Panels show the four hedonic specifications.

C Land Value Estimation

Baseline: GAM-Predicted Land Values from Vacant Land Transactions

Transaction data and outlier treatment. The analysis sample consists of arms-length vacant land transactions recorded in the MAPC region from 2010 to 2019, merged with parcel characteristics from ATTOM assessor records. I apply two sequential filters before estimation. First, I remove observations on substantive grounds: transfers below $1,000 (nominal), parcels with missing or non-positive lot size, prices above $1,000 per square foot, and non-arms-length document types (foreclosure, quitclaim, and trustee deeds). Second, I estimate a preliminary OLS regression of log price per square foot on zoning district fixed effects, log lot size, and year fixed effects, and discard observations whose residuals exceed three standard deviations. This filter removes statistical outliers while retaining genuinely high-value transactions in expensive zoning districts, dropping 13 transactions.

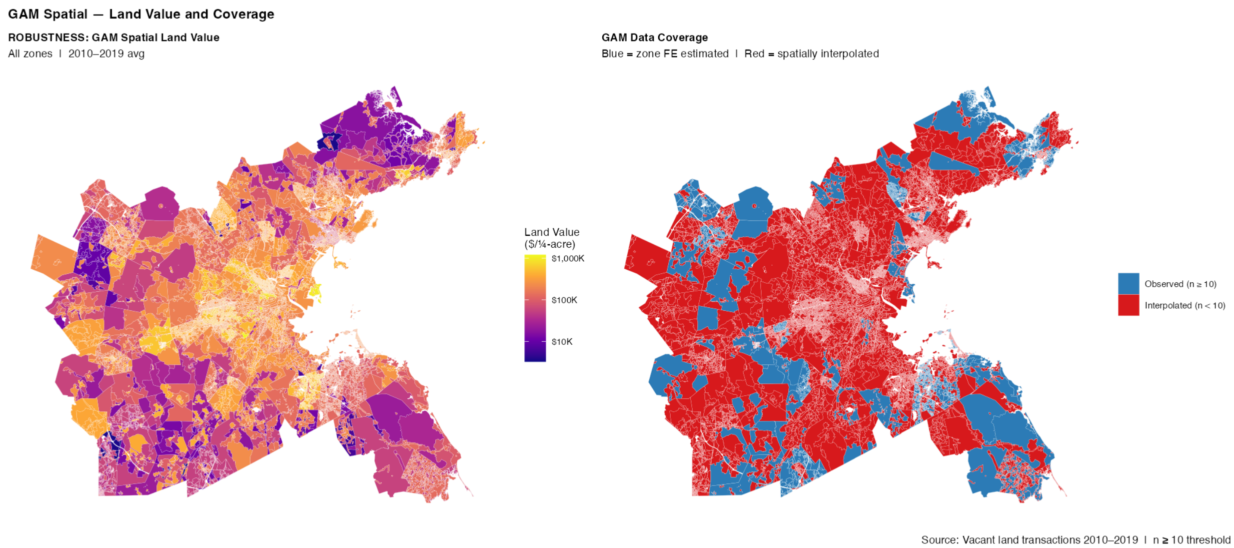
The filtered sample contains 2,067 transactions across 461 of 1,775 zoning districts in the MAPC region. Vacant land transactions are sparse: only 56 districts (3.2% of all districts) have at least ten observations and contribute directly to GAM estimation, while the remaining 405 districts with fewer than ten transactions and 1,314 districts with zero transactions are spatially interpolated from the fitted surface. This sparsity motivates the continuous spatial smooth specification rather than zoning-district fixed effects, which would leave the majority of districts without predicted values.

Model specification. I estimate the following generalized additive model by restricted maximum likelihood (REML) using the `mgcv` package in R:

$$\log(\text{price per sq. ft.})_{ij} = f(x_j, y_j) + g(\text{lot size}_{ij}) + \alpha_{\text{muni}(j)} + \gamma_t + \varepsilon_{ij}, \tag{1}$$

where i indexes transactions, j indexes zoning districts, and t indexes years. $f(\cdot)$ is a thin-plate regression spline in zoning district centroid coordinates (x_j, y_j), with smoothness selected automatically by REML. $g(\cdot)$ is a smooth function of lot size. $\alpha_{\text{muni}(j)}$ are municipality fixed effects and γ_t are year fixed effects. Predicted values are generated at each zoning district centroid for each sample year and averaged to form the zoning district-level land value measure used in the structural model. Zoning districts with fewer than ten observed transactions are classified as spatially interpolated; their predicted values are derived entirely from the fitted smooth surface. Figure A.12 shows GAM-predicted land values and data coverage across zoning districts.

Figure A.12 — GAM-Predicted Land Value and Data Coverage



Note. The left panel shows GAM-predicted zoning district-level land value averaged over 2010–2019, converted to a $\frac{1}{4}$-acre basis. The right panel shows data coverage: blue zoning districts have at least ten observed vacant land transactions and contribute directly to GAM estimation; red zoning districts are spatially interpolated from the fitted surface. Source: vacant land transactions, MAPC region, 2010–2019.

Comparison with FHFA Land Price Estimates

As an external benchmark, I compare the GAM-based land values with the Federal Housing Finance Agency (FHFA) experimental land price dataset (Davis et al., 2021), which estimates the value of land underlying single-family structures from cost-approach appraisals. The FHFA dataset is constructed from the Uniform Appraisal Dataset (UAD), covering single-family properties financed through Fannie Mae and Freddie Mac. The cost approach separates total appraised value into a structure replacement cost and a residual land component. To reduce anchoring to tax-assessed values, the FHFA procedure excludes appraisals within 2% of the assessed value for either the land component or the total value, yielding approximately 7.9 million valid appraisals nationwide. All values are nominal. The standardized land value series normalizes each estimate to a $\frac{1}{4}$-acre lot using a hedonic regression of log land value on log lot size, estimated separately by county with ZIP code and year fixed effects.

I use version 3.0 (October 2020), which covers 2012–2019 at the census tract level across 48 states.[1] Because FHFA estimates are reported at the census tract level while my analysis operates

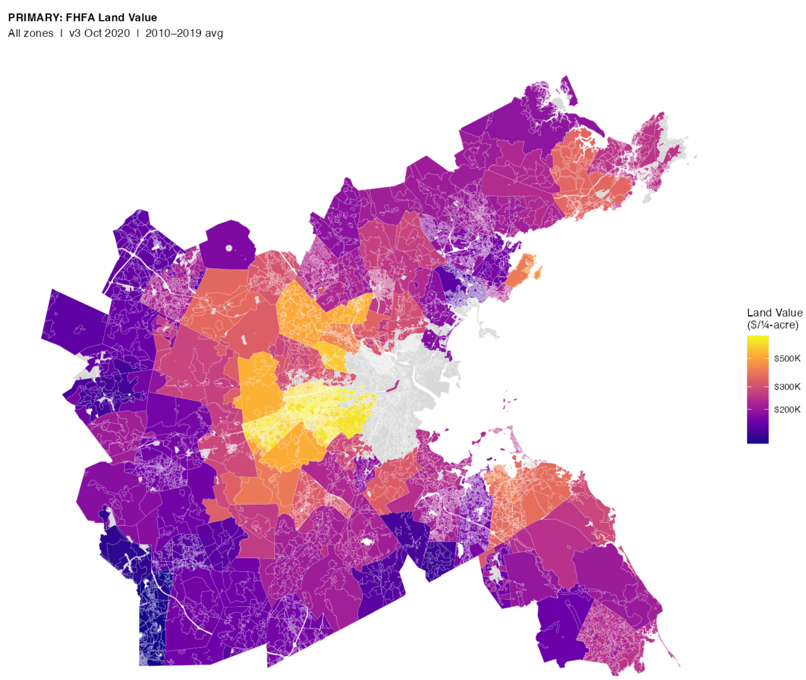
1. Version 4.0 (June 2024) extends coverage through 2022, adds two states, and corrects a log-to-level transformation error present in earlier releases that caused land values to be understated by 1–3%. I use version 3.0 as my

at the zoning district level, I aggregate tract values to zoning districts using area-weighted averaging. For each zoning district z I compute

$$\hat{v}_z = \sum_{t \in \mathcal{T}(z)} w_{zt} \cdot v_t, \qquad w_{zt} = \frac{|A_z \cap A_t|}{|A_z|}, \tag{2}$$

where v_t is the FHFA land value for tract t, A_z and A_t are the geometries of zoning district z and tract t, and $\mathcal{T}(z)$ is the set of tracts intersecting z. Weights sum to one within each zoning district by construction. The aggregation uses 2010 TIGER/Line boundaries, which match the tract definitions used by FHFA throughout both vintages. Figure A.13 shows the resulting zoning district-level estimates averaged over 2012–2019.

Figure A.13 — FHFA Land Value



Note. This figure shows FHFA land value estimates aggregated to MAPC zoning districts, averaged over 2012–2019. Values represent the estimated price of a standardized $\frac{1}{4}$-acre lot. Aggregation uses area-weighted averaging from 2010 census tracts (equation 2). Source: FHFA Land Price Dataset v3.0 (October 2020).

Figure A.14 compares the two measures for zoning districts with at least ten transactions. The two series are positively correlated but exhibit a systematic level difference: FHFA estimates exceed GAM-predicted values on average, and the log-log slope (with FHFA on the vertical axis)

primary vintage given its closer alignment with my 2010–2019 sample period.

is below one, indicating that the FHFA measure compresses the cross-sectional range of land values relative to the GAM.

The FHFA dataset derives land value residually from cost-approach appraisals: for each property, the appraiser estimates the depreciated replacement cost of the structure, and land value is computed as the difference between appraised house value and structure cost. The GAM measure, by contrast, reflects prices actually paid in arms-length transactions for vacant parcels. These two concepts can diverge for several reasons: vacant parcels may trade at a discount relative to land under improved structures if they lack infrastructure or entitlements, or at a premium in dense urban areas where they embed development option value; and the FHFA sample is restricted to GSE-financed properties, which excludes the highest-value homes above conforming loan limits.

Figure A.14 — FHFA vs. GAM-Predicted Land Value

Note. This figure compares FHFA and GAM-predicted land values at the zoning district level for zoning districts with at least ten vacant land transactions over 2010–2019. The left panel uses the GAM zone fixed effects specification; the right panel uses the GAM spatial smooth specification. Both axes are on a log scale. The dashed line is the 45-degree reference; the red line is an OLS fit with a 95% confidence band.

D Model Estimation with Explicit Zoning Controls

The baseline model represents explicit zoning through a single dimension, the minimum lot size, and attributes the remaining gap between predicted profitability and observed development to the municipality-specific implicit costs $\omega_{m,1}$ and $\omega_{m,2}$. A natural concern is that $\omega_{m,1}$ absorbs other explicit zoning rules omitted from the baseline. This appendix re-estimates the model with additional district-level zoning controls and shows that the estimated implicit development costs remain largely unchanged.

Augmented model. I add a vector of district-level zoning characteristics Z_i to the development payoff in Equation (1):

$$\Pi_{ik} = \frac{P_{ik} - C_{ik} - \omega_{m(i),1} - \omega_{m(i),2} \cdot \mathbb{1}(x_{1,k} < \bar{x}_{1,i}) - Z_i'\psi - \gamma_0 \cdot s_i \cdot \mathbb{1}(\text{improved}_i)}{x_{1,k}} + \varepsilon_{ik}. \quad (3)$$

The coefficient vector ψ is common across municipalities. It must be pooled rather than municipality-specific because any municipality-level zoning measure is collinear with the free constant $\omega_{m,1}$; ψ is therefore identified only from variation in Z_i across zoning districts within municipalities. The model is estimated jointly by stacking the municipality likelihoods, so the municipality-specific $\{\omega_{m,1}, \omega_{m,2}\}$ and the single ψ are recovered in one maximization.

Control set. Z_i contains the maximum floor-area ratio, the maximum building height, the maximum lot coverage, and an indicator for whether multifamily housing is permitted by right. Continuous controls use the effective (harmonized) values from the MAPC Zoning Atlas, standardized across regulated districts; districts where a regulation is absent are set to the mean and flagged with a separate existence indicator, so "unregulated" enters as its own category rather than as a missing value.

I exclude the maximum dwelling units per acre and the maximum number of dwelling units. For single-family parcels these are mechanically tied to the minimum lot size already in the model. Density is approximately one unit per lot, so the permitted density in dwelling units per acre is roughly $43{,}560/\text{MLS}$. Including them would enter the same constraint twice and, more importantly, would be internally inconsistent in the counterfactuals, where removing the minimum lot size requirement would by construction change a density cap denominated in units per acre, so it cannot be held fixed as a separate control.

Results. Table 6 reports the resulting regulatory costs in the row "Add explicit zoning controls." The median $\omega_{m,1}$ moves only from \$117k to \$114k and the median $\omega_{m,2}$ from \$45k to \$39k; across municipalities the estimated $\omega_{m,1}$ correlates 0.94 with the baseline and $\omega_{m,2}$ correlates 0.80. The implicit cost is thus not a mechanical artifact of omitted explicit zoning. Controlling directly for floor-area, height, coverage, and multifamily rules leaves the municipality-specific costs largely unchanged. As noted in the main text, the mean and standard deviation of $\omega_{m,2}$ rise under the controls, driven by a small number of municipalities where $\omega_{m,2}$ is weakly identified, while the median and the cross-specification correlation are the stable statistics. The estimated control coefficients ψ are reported in Table A.6, although they are weakly identified (footnote 6) and are not interpreted individually.

Table A.6 — Estimated Coefficients on Explicit Zoning Controls

Control (Z)	ψ (\$000s)
Max. floor-area ratio (per SD)	78.8
FAR regulated (indicator)	−34.6
Max. height (per SD)	12.1
Height regulated (indicator)	−70.4
Max. lot coverage (per SD)	8.9
Lot coverage regulated (indicator)	65.5
Multifamily allowed by right (indicator)	70.8

Note. The table reports pooled coefficients ψ on the district-level zoning controls Z_i in Equation (3), in thousands of dollars per project, estimated jointly with the municipality-specific costs. Continuous controls are standardized over regulated districts, so their coefficients give the dollar change per one standard deviation, and each is paired with an indicator for whether the regulation exists in the district. These coefficients are weakly identified and are not interpreted individually (footnote 6); the offsetting signs within each pair reflect the near-collinearity of the standardized value and its existence indicator when a regulation is recorded in few districts.