Elasticity or Shift? Why Housing Becomes Unaffordable*

John D. Baker† Shahar Rotberg‡

September 25, 2026

Abstract

Why do house prices grow faster in denser places? Common answer: low elasticity. But prices cannot grow if supply shifts one-for-one with demand, however inelastic. We show that supply shifts, not the slope, drive most of the price-growth gap. To do so, we use U.S. census-tract data to recover demand and supply shifts, and find that supply's absorption of demand pressure falls sharply with density, a pattern that holds in the recovered shifts, in raw housing-unit counts, and under predicted demand shift-shares no locality chose. A theory of the buildable cap rationalizes both: the room left under the cap determines the elasticity, and the cap growth affects the shift. We estimate that dense places relax their caps roughly ten times more slowly than sparse ones. Small increases in the cap generate the supply shifts the data require, and the caps we recover are supported by independent regulation measures.

Keywords: Housing supply, density, supply elasticity, affordability, zoning.
JEL Codes: R31, R14, R52, H70.

1 Introduction

Housing affordability has become a central economic and political concern across the United States. Between 1999 and 2016, quality-adjusted real house prices in the densest tracts rose by about 108 percent, while prices in the least dense tracts rose by about 61 percent. What drives this disparity?

The standard economic explanation is that the gap is driven by more inelastic supply in denser locations: dense places sit on a steep supply curve, so any given demand shock moves prices more than quantities (Glaeser et al., 2005; Saiz, 2010; Gyourko and Molloy, 2015). However, if supply shifts one-for-one with demand, prices do not grow, no matter how steep the curve is. Only a supply that persistently falls short

*The opinions expressed herein are those of the authors and do not necessarily represent the opinions of the Canada Mortgage and Housing Corporation (CMHC). We thank Lu Han, Tom Davidoff, John Mondragon and Efraim Benmelech for helpful comments and insights. We used Claude and refine.ink in the creation of this draft. The replication package is available at https://github.com/srotberg/elasticity_or_shift_why_housing_becomes_unaffordable.

†Canada Mortgage and Housing Corporation. Email: jbaker@cmhc-schl.gc.ca.
‡Canada Mortgage and Housing Corporation. Email: shahar.rotberg@gmail.com; web: www.shaharrotberg.com.

of demand can explain prices growing year after year. Whether dense places are unaffordable because their supply curves are steep or because those curves shift out too slowly is therefore an important and underexplored distinction. Our first goal is to separate the two in the data.

Once separated, we ask two additional questions. First, does supply in general shift one-for-one with demand, so that building keeps pace and prices are stable? Second, does supply shift differently in denser places? Denser environments have scarcer land, more complex regulation, and more contested construction, so supply may respond less there.

Density is a key object because it is where the policy debate lives: the contested zoning fights, the heritage designations that freeze what may be rebuilt, the proposals for by-right development and transit-oriented upzoning, and the political economy of neighborhood opposition all play out in already built-up places. If the supply response is systematically weaker exactly there, then demand growth in the places people most want to live is converted into price growth rather than into homes. The affordability problem then becomes self-reinforcing: the more desirable and built-out a location becomes, the less its supply responds to the next increment of demand.

To separate the shift from the slope, and to answer the two questions above, we turn to the data in two empirical steps. First, we use a structural model of housing demand and supply to recover model-implied demand and supply shifts from observed prices and quantities, at the U.S. census-tract level, from 1999 to 2016. The goal of this step is to produce economically interpretable measures of demand pressure and supply response that are consistent with the data. Second, we examine how the recovered supply shifts vary with demand shifts and with local density, through an interaction that tests whether density amplifies or dampens the responsiveness of supply. Throughout, density serves as a sufficient statistic for constraints that are only partially observable at the tract level and are themselves endogenous to prices.

We do not estimate the two elasticities ourselves but draw them from existing work. We take the supply elasticity from the neighborhood-level estimates of Baum-Snow and Han (2024), so each census tract carries its own rather than an assumed one, and the demand elasticity from the household literature that estimates it directly (Hanushek and Quigley, 1980; Polinsky and Ellwood, 1979; Mayo, 1981; Albouy et al., 2016). All of our results are robust to these choices: flattening the supply elasticity gradient entirely, rescaling it to three times its estimated slope, and perturbing every tract's elasticity independently within a wide band. Results continue to hold when varying the demand elasticity across a range that brackets that literature's estimates and letting it differ by tract with income and with a tract's measure of substitutability (Appendices A.5 and A.6).

We obtain three main results. First, housing supply only partially tracks demand. Applied to the average tract's price and stock growth, the recovery puts the supply shift at roughly 0.26 of the demand shift, so that adjustment occurs through a combination of new building and rising prices rather than through new building alone.

Second, this responsiveness falls sharply with density. The marginal supply response to demand declines

by about 68 percent from the least to the most dense decile, from about 0.261 to about 0.083, with the median decile at about 0.148.

Holding demand growth and the supply elasticity fixed across the comparison, these absorption gaps compound into divergent price growth. Fed through the model, the densest decile's real house prices grow about 0.24 percentage points per year faster than the median decile and about 0.65 points faster than the least dense, leading to prices that rise about 1.08 and 1.31 times as much over the sample, respectively. Combined with the output elasticity of Hsieh and Moretti (2019), the densest-to-median gap alone implies forgone output of roughly $64 to $127 billion per year, about 0.35 to 0.7 percent of GDP, which we offer as an illustrative magnitude.

Because the supply elasticity also falls with density, we separate the elasticity and shift channels by counterfactual. In equilibrium, price growth is unabsorbed demand growth divided by the sum of the absolute price elasticities of supply and demand. The shift therefore sets the numerator, how much demand goes unabsorbed, and the slope sets the denominator, how strongly that remainder is capitalized into price. Had supply shifted one-for-one with demand, the numerator would be zero and prices would not have grown at any elasticity.

Our third result follows: the slope accounts for only about 29 percent of the density gradient in price growth, with the remainder attributed to the shift (Section 5.2). Moreover, doubling the densest decile's supply elasticity lowers its price growth by only about a quarter as much as doubling its absorption of demand does, and under every calibration we examine the elasticity remains the weaker lever (Section 8.4).

One concern is that localities may shape not only how much they build but also how much demand they attract. To address this, we re-run the estimation with demand measured by a Bartik shift-share. This method estimates the demand growth implied by each county's pre-sample industry mix and national industry fates, which no locality could have chosen and which uses no housing data at all. It delivers an even stronger density gradient, and a placebo supports the design: the same instrument produces no density gradient in the decade before our sample, so what we measure is not a pre-existing tendency of dense tracts to build less (Section 5.3).

Applied to raw housing-unit counts, the same instrument shows the gradient outside our model altogether: places with the same predicted demand growth build less housing where density is higher (Section 5.3).

Lastly, Section 6 develops a theory of the buildable cap that rationalizes our results. The elasticity and the shift are two properties of a single object, the cap: how much room it still leaves determines the supply elasticity, and how fast it rises affects the supply shift. The theory summarizes that rise in the cap by a cap-relaxation function, which we then estimate. This yields a further result: caps in the densest decile are relaxed at roughly one-tenth the pace of those in the least dense deciles, which raise theirs by about 0.1 percent to 0.2 percent a year.

Where the cap can be read directly off administrative records, it matches: New York City's lot-level

zoning files show the median tract's allowed residential capacity unchanged from 2002 to 2016 (Section 7.2). Regulation measured independently, in the Wharton land-use survey, corroborates the pattern. Following the same communities across the survey's two waves, we find that regulation loosened less, or tightened more, where density was high than where it was low, with the shortfall concentrated among dense metropolitan areas facing strong demand, which is where standing close to the cap should bind.

The paper proceeds as follows. Section 2 places the paper in the literature. Sections 3 to 5 set out the recovery, the data, and the estimates. Section 6 develops the theory that rationalizes them and estimates the cap-relaxation function. Section 7 confronts the recovered caps with regulation measured from outside the model, Section 8 reports the robustness analyses, and Section 9 concludes.

2 Related literature

Our work relates to a large literature on the elasticity of housing supply. Glaeser et al. (2005) and Saiz (2010) document substantial heterogeneity in supply elasticities across U.S. metropolitan areas, emphasizing geographic constraints and land-use regulation. Green et al. (2005) provide some of the earliest metropolitan-specific estimates and show that they vary systematically with density and regulation, an antecedent of the neighborhood-level gradient we exploit. Gyourko et al. (2008) and Gyourko and Molloy (2015) document the importance of regulatory restrictiveness and conclude that regulation, rather than geography alone, is the binding margin in the most expensive markets, and Hilber and Vermeulen (2016) find the same in English planning-refusal data. Glaeser and Gyourko (2018) show that the wedge between house prices and production costs has widened precisely in dense coastal markets, the places our estimates identify as least responsive. Paciorek (2013) embeds these constraints in a dynamic model in which regulation raises price volatility by slowing the supply response, a mechanism close in spirit to the shift we measure.

The elasticity has proved its worth beyond the housing literature: Mian and Sufi (2011) and Mian et al. (2013) instrument house price growth with Saiz's measure to identify how home-equity borrowing affects short-run changes in consumption. Over their short horizons the shift is largely absent: a supply shift is the buildable cap moving, which means regulation changing, and that is a long process that takes many years. What remains is building within limits already granted, and that is what the elasticity measures. Over such horizons the elasticity is essentially what separates one place's short-run price fluctuations from another's. Our aim is to understand long-run price trends rather than short-run fluctuations, and along these horizons regulations and supply have enough time to respond to changes in demand and the economic environment.

Taken together, the current literature characterizes supply constraints through their effect on movements along a given supply curve, that is, through the elasticity. We depart from it by measuring shifts of the curve itself, recovering model-implied demand and supply shifts from equilibrium outcomes and asking how much of a demand increase is accommodated by building rather than by higher prices. We build most directly on Baum-Snow and Han (2024), whose estimated elasticities decline with density and rise with land

availability and looser regulation. Theirs are the first elasticities measured *within* cities rather than across them: where the literature above compares one metropolitan area to another, their estimates let us give each census tract its own elasticity. Heterogeneity in measured elasticities, however, matters for long-run affordability only insofar as supply shifts fall short of demand.

Closest to our mechanism is Rollet (2025), who assembles a parcel-level panel of roughly 833,000 New York City lots from 2004 to 2022 and estimates a dynamic model of redevelopment. Two of his findings are ours seen from the ground. Redevelopment responds to the room a parcel has left: a lot needs about two additional points of permitted floor-area ratio before it has even a fifteen percent chance of being rebuilt over eighteen years, which is direct micro evidence for the headroom our elasticity measures. And the fixed cost of redevelopment rises steeply with the size of the building already standing, so relaxing the cap where much already stands buys less building than the same relaxation elsewhere. That is the force our extension in Section 6.6 puts into the cost function directly.

A separate strand asks not how large supply constraints are but why they exist, and why they bind hardest where the pressure to build is greatest. Favilukis and Song (2023) give a political-economy answer: when zoning authority is fragmented across many small jurisdictions, local homeowners impose stricter regulation to limit congestion and other neighborhood externalities, and because each jurisdiction is small it does not internalize the cost to metro-wide affordability. They show that more fragmented metropolitan areas have more restrictive zoning and higher housing costs, identified from fragmentation shaped more than a century ago. Ortalo-Magné and Prat (2014) supply the complementary voting logic: because homeowners' wealth is tied to house values, political support for restricting supply is strongest exactly where prices are already high. Einstein et al. (2019) find the same forces at work in local planning meetings, hand-coding their minutes to show that participants are unrepresentative of their communities and overwhelmingly opposed to new construction, which translates into fewer permits and into the delays and downsizing that make projects unviable.

These are not the geographic limits often invoked to explain high prices but rules and processes that could be changed. All three mechanisms predict that the willingness to permit falls precisely where demand presses hardest. Our finding that the supply shift weakens with density is the equilibrium footprint of that logic, measured in the supply response itself rather than in the stringency of regulation.

Demand, meanwhile, concentrates where supply responds least: Diamond (2016) shows that high-skill workers increasingly sort into dense, high-amenity metropolitan areas and that amenities respond endogenously to their arrival. The aggregate stakes of weaker supply shifts there are set out by Hsieh and Moretti (2019), who argue that stringent constraints in high-productivity cities such as New York City and San Francisco prevent workers from moving to where their productivity would be highest, misallocating labor across space. In their corrected estimates, relaxing land-use restrictions in those cities to the median U.S. level would raise aggregate output by several percent. Section 5.2 puts a magnitude on this channel by combining our density gradient with their output elasticity. Ganong and Shoag (2017) trace the same forces over the

long run, showing that incomes in poorer states had for a century been catching up with incomes in richer ones, and that this catch-up stalled as land-use constraints tightened and high housing costs priced out the low-skill migrants for whom the move used to pay. In aggregate models these constraints enter as regulatory wedges: Herkenhoff et al. (2018) show that the tightening of land-use restrictions across U.S. states since the 1950s can account for a sizable share of the slowdown in aggregate output and labor reallocation. The cap-relaxation rate we recover is the local margin such wedges aggregate over.

A final strand is one we use rather than extend. Because our framework recovers demand growth as a residual, it needs an external yardstick, and the literature estimating how far house prices move with mortgage rates supplies one. Damen and Schildermans (2022) exploit a change in bank capital requirements and estimate a semi-elasticity of real house prices to the mortgage rate of about 12.3 percent per percentage point. Laufer and Tzur-Ilan (2021) reach a comparable figure from a different institutional experiment, combining the price effect of a 2010 Israeli macroprudential rule that raised risk weights on high loan-to-value mortgages with the effect of the same rule on the rates borrowers paid, and obtain a range of 6 to 10. Designs built on the conforming loan limit, such as DeFusco and Paciorek (2017), report smaller numbers, but they identify the interest-rate elasticity of *mortgage demand* rather than of house prices, so they answer a different question. Section 4.1 runs our recovered demand growth back through the experiment these studies conduct and shows that it implies a semi-elasticity of the size the first two studies report.

Lastly, our study speaks to a recent challenge to the supply-constraint view. Louie et al. (2026) show that income growth predicts essentially the same house-price and quantity growth across U.S. metropolitan areas regardless of their measured supply elasticity, and conclude that measured supply constraints do not explain the growth of housing prices and quantities. Our results suggest a different reading. First, we show that their price null result is a matter of scale: when we run the same regression across census tracts, measured elasticity does predict how strongly income growth passes into prices, so the relationship is present neighborhood by neighborhood but masked when aggregated to the city level. The slope is not missing from the data; it is harder to see at the scale they measure it, because a metropolitan elasticity averages over neighborhoods standing close to their caps and neighborhoods with ample room. Second, we show that the slope is not irrelevant but is the smaller channel: in our shift-elasticity composition it carries about 29 percent of the density gap in price growth, against 71 for the shift. Their own analysis points to the supply shift as the explanation, but they reject it, because in their model a shift can only mean implausible productivity gains; Appendix A.13 shows the required shift can instead be rationalized as ordinary movement of the buildable cap, which independently measured regulation supports (Section 7).

Relative to all of these strands, our contribution is to measure the object they leave implicit. The elasticity literature measures the slope of the supply curve; the political-economy literature explains why the constraints behind it exist; the misallocation literature counts the aggregate cost. What none of them measures is the outward movement of the supply curve itself, and that movement, more than the slope, is what accumulates into long-run unaffordability. We recover it neighborhood by neighborhood, show that it falls

systematically with density, and trace the consequences for price growth. In doing so we also provide the regulation literature with a complement to its survey-based indices: the cap-relaxation rate recovered in Section 6.5 is a revealed-behavior measure of the *pace* of regulatory change, available for every census tract, where the indices measure the level of stringency for the communities they survey.

3 Methodology

3.1 Overview

To measure how much housing supply responds to demand, and how that response varies with local density, we adopt a two-step approach. First, we use a structural model of demand and supply to recover model-implied measures of demand and supply shifts from observed prices and quantities. Second, we examine how supply shifts vary with demand shifts and local density, and in particular whether denser areas exhibit systematically weaker supply shifts.

3.2 The supply and demand model

Our demand and supply functions take the following forms:

$$D_t = A_1 \, \varphi_t \, p_t^{\varepsilon_d}, \tag{1}$$

$$S_t = A_2 \, \psi_t \, p_{t-m}^{\varepsilon_s(\text{Den}_{t-m})}. \tag{2}$$

In (1) the price is raised to the power of the demand elasticity ε_d, and in (2) it is raised to the supply elasticity ε_s. The supply elasticity is a function of local density in period $t-m$, $\varepsilon_s(\text{Den}_{t-m})$, where m is the construction lag introduced below; as the Baum-Snow and Han (2024) estimates show, denser locations have a lower supply elasticity and hence a steeper supply curve.

The terms A_1 and A_2 are scalars, and φ_t and ψ_t are unknown demand and supply shifters normalized to one in the base period. The lag m on the price in (2) reflects the idea that construction takes time to complete. A developer observes the price, decides to build, and delivers the units m years later, so building completed in year t is a function of the price in year $t-m$; we set m in Section 4. The lag is a time-to-build delay, not a coarsening of the data: the flow is annual, and only the first m years of the panel are lost to forming the lag.[1]

Because we observe the equilibrium price and quantity each period, and the shifters equal one in the base period, we recover the scalars from the base period and then the shifters period by period. Because the base-period supply scalar A_2 requires a price m years earlier, the base period is not the panel's first

[1] Responding to the lagged price is the myopic reading of time to build: the developer who decides at $t-m$ acts on the price prevailing at that time. The opposite bound is perfect foresight, under which the developer forecasts the delivery-date price exactly, so that supply responds to the current price. Re-running the entire recovery under that bound gives a density interaction of -0.098 ($t=-33$) against the baseline -0.036 (Section 5), so the gradient survives, and is in fact steeper, at either extreme of how developers form expectations.

year t_0 but $t_0 + m$, the first year for which both a current and a lagged price exist. The first m years of the panel supply lagged prices but do not themselves receive recovered shifters. Writing $b \equiv t_0 + m$ for the base period, so that $\varphi_b = \psi_b = 1$, the base-period demand and supply relations pin down

$$A_1 = \frac{Q_b}{p_b^{\varepsilon_d}}, \qquad A_2 = \frac{Q_b}{p_{t_0}^{\varepsilon_s(\mathrm{Den}_{t_0})}}, \tag{3}$$

where the base-period supply responds to the price m years earlier, which is the price in the panel's first year t_0.

Demand shifter recovery. To obtain the demand shift in later periods, we compare the observed stock to the quantity the base-period demand function would predict at the prevailing price, which in logs gives

$$\log \varphi_t = \log(Q_t / Q_b) - \varepsilon_d \log(p_t / p_b). \tag{4}$$

Because φ equals one in the base period, φ_t measures demand in period t relative to that base: a value of 1.2 means demand has grown by twenty percent since. We therefore write $\varphi_t = 1 + \Delta D_t$, where ΔD_t is that proportional growth, so that $\log \varphi_t = \log(1 + \Delta D_t)$.

Supply shifter recovery. To obtain the supply shift, we equate quantity demanded to quantity supplied at the observed prices in periods t and $t - m$. Taking logs, the supply shifter satisfies the intermediate expression

$$\log \psi_t = \log(A_1 / A_2) + \varepsilon_d \log p_t - \varepsilon_s(\mathrm{Den}_{t-m}) \log p_{t-m} + \log \varphi_t.$$

When we substitute the demand shifter from Equation (4), the period-t price terms cancel, since they enter the demand shift and the supply shift with opposite signs. What remains is the recovered log supply shift as a function of period t's equilibrium quantity, the supply elasticity (itself a function of density in period $t - m$), period $t - m$'s price, and a constant,

$$\log \psi_t = \log Q_t - \varepsilon_s(\mathrm{Den}_{t-m}) \log p_{t-m} - \log A_2. \tag{5}$$

As with demand, ψ equals one in the base period, so we write $\psi_t = 1 + \Delta S_t$, where ΔS_t is the proportional growth in supply since that base and $\log \psi_t = \log(1 + \Delta S_t)$. The supply shift is therefore a percentage change throughout: wherever we say a location's supply shifts by more or by less, we mean in percentage terms, not in units built.

3.3 Calibrating the supply elasticity from Baum-Snow and Han

A central input into the recovery is the supply elasticity. The supply elasticity determines how much of an observed quantity change the model credits to movement along the supply curve, and hence how much remains to be attributed to the shift. Rather than choosing a value ourselves, we discipline it with the most directly relevant external evidence available, the neighborhood-level estimates of Baum-Snow and Han (2024). They estimate housing-unit supply elasticities at the census-tract level and show how these vary with local determinants of supply. Their object is the response of the *total* unit stock over 2000 to 2010, from all margins of adjustment: new construction, teardowns, and reconfiguration. Their unit elasticities average about 0.39 across tracts and rise with distance from the central business district, the availability of developable land, and looser regulation.

One adjustment is required before these values enter our model. The Baum-Snow and Han elasticity is a *cumulative decadal* object: it measures how much the stock changes over roughly ten years in response to a price change, summing the construction that completes across the entire period. Our model, by contrast, is annual: it relates each year's flow of completions to the price m years earlier. Because ten years of cumulative construction is the sum of ten annual flows, the elasticity governing a single year's flow is one-tenth of the decadal cumulative elasticity. We therefore divide the Baum-Snow and Han elasticities by ten to obtain the annual-flow elasticity our model requires. This brings the average to about 0.039 across tracts, with a median of 0.036.

Reassuringly, this is the same order of magnitude as the elasticities used in related work on annual data (Mayer and Somerville, 2000; Paciorek, 2013), so the conversion aligns the external evidence with the object the model needs rather than importing a mismatched horizon.[2]

Rather than assuming a functional form for how the supply elasticity varies with density, we estimate it from the Baum-Snow and Han data. Their tract-level elasticities are strongly and monotonically related to density. Fitting a log-linear relationship across tracts at their density in period $t - m$ gives

$$\varepsilon_s = 1.25 - 0.125 \log(\text{Den}_{t-m}), \tag{6}$$

with log density alone explaining more than half of the cross-tract variation in elasticity ($R^2 \approx 0.54$).

Figure 1 plots this relationship in Baum-Snow and Han's *decadal* elasticities: it falls from about 0.74 in the least dense decile to about 0.13 in the densest. Equation (6) measures directly in the data what Section 6 derives in theory: that the elasticity declines with density. Because the gradient is a function of density, it also imputes an elasticity to every tract, including the small share for which Baum-Snow and Han could not compute a value, purely from observable density.[3]

[2]The divide-by-ten is a first-order approximation, so we return to it in the robustness analysis: as Appendix A.8 shows, our central result is unaffected across a wide range of alternative conversion factors.

[3]Because a fitted line extrapolates beyond the support of the estimates at extreme densities, we bound the imputed elasticity: below at zero, since a negative supply elasticity is not economically meaningful and the few negative values among the Baum-Snow and Han

We do not fix the elasticity of a tract over time. Instead, we let each tract's elasticity move along the empirical gradient (6) as its own density evolves, evaluating (6) at the tract's lagged density in each period. A neighborhood therefore becomes less supply-elastic as it builds out. This lets the density mechanism operate within a tract over time, not only across tracts.[4]

3.4 Our econometric model

Given the recovered shifts, we estimate how the supply shift varies with the demand shift and density, writing the specification so that the demand shift enters first:

$$\log \psi_{i,t} = \beta_0 + \beta_1 \log \varphi_{i,t} + \beta_2 \left(\log \text{Den}_{i,t-m} \times \log \varphi_{i,t} \right) + \beta_3 \log \text{Den}_{i,t-m} + \epsilon_{i,t}. \tag{7}$$

Density enters at $t - m$, the same predetermined density that governs the supply elasticity in (2). The coefficient β_1 captures the baseline responsiveness of supply to demand: a value of one means supply shifts one-for-one with demand, and a value below one means supply shifts less than demand. The coefficient β_2 captures how density modifies the responsiveness of the shift, and a negative value indicates that density dampens it. The coefficient β_3 measures the effect of density on supply shifts in the absence of demand pressure.

Our preferred specification omits tract and year fixed effects, but our results are robust to including them (Appendix A.3 and Section 8.1). The substantive reason to omit tract fixed effects is that the structural recovery already absorbs systematic differences across locations. Because the demand and supply shifters are recovered from each tract's own observed prices and quantities, any systematic feature of a location, its geography, regulatory environment, amenities, or position in the urban hierarchy, is already embedded in the recovered shifts.

We omit year effects in our preferred specification because our object of interest is the general-equilibrium relationship between demand pressure and supply response, including the common, system-wide growth in demand over this period, which we mean to study rather than treat as a confound.[5]

3.5 How the elasticities affect the estimates

Because the recovered shifts are constructed for given elasticities, it is essential to separate their effect on magnitude from their effect on sign. The recovery shows where each elasticity enters, which tells us where to look.

(2024) tract estimates are noise around zero in built-out tracts, and above at the largest elasticity they estimate, about 1.17 in decadal units. The bounds bind for fewer than 0.2 percent of tracts.

[4]The gradient is not what produces the density dampening, as Section 3.5 explains and Section 8 verifies: flattening it entirely, so that every tract carries the same elasticity, leaves the interaction at -0.031 against a baseline of -0.036, and holding each tract's elasticity fixed at its 2000 value leaves it negative and significant as well.

[5]Appendix A.2 re-estimates the baseline with standard errors clustered by tract, by year, and two-way, and Appendix A.3 adds metro and metro $\times$ year fixed effects, which compare dense and sparse tracts within the same city and so net out anything specific to a metropolitan area or to a metro-year. Our results are robust to these adjustments.

First, the demand elasticity cancels out of the recovered supply shift, as shown above. It enters the regression only through the demand-shift regressor, $\log \varphi = \log Q_t - \varepsilon_d \log p_t$. Changing ε_d does not simply rescale this regressor. A change of $\Delta\varepsilon_d$ subtracts $\Delta\varepsilon_d \log p_t$ from it, and because cumulative price growth differs across tracts, and systematically with density, the regressor changes by different proportions in different tracts. A pure rescaling could move only the magnitude of the coefficients; a change of this kind has no such algebraic guarantee, so whether the sign survives is an empirical question. In practice only the magnitude moves: the sign holds across a range of demand elasticities that brackets the literature's estimates (Appendix A.5), and Section 5.3 replaces the regressor with predicted demand, which involves no demand elasticity at all. A uniform change in the level of the supply elasticity likewise leaves the sign intact across a wide range of conversion factors (Appendix A.8).

Second, the input with the most direct potential to affect the sign of the interaction is the relationship between the supply elasticity and density. Because the recovered supply shift contains the term $-\varepsilon_s(\text{Den}_{t-m}) \log p_{t-m}$, letting the supply elasticity vary with density injects a density-correlated component into the dependent variable. Whether that component pushes the interaction up or down is not something the algebra settles, so this is an empirical question. Section 8.2 answers it by rescaling the gradient from zero to three times its estimated slope: the interaction stays between -0.029 and -0.036 throughout, so the gradient affects the estimate only marginally and does not generate it.

3.6 Endogeneity

A careful reader may worry that localities choose more than their supply policies: they may also influence their demand growth directly, and if the two choices are correlated, our coefficients are biased. To address this concern, in Section 5.3 we build a Bartik shift-share, which is the demand growth predicted by each county's pre-sample industry composition, and we use it both in place of the recovered demand shift and as an instrument for it; the gradient survives and comes out stronger. Section 7 then repeats the comparison with measured regulation, again using predicted demand in place of the recovered demand. If the density gradient were an artifact of localities choosing their own demand, replacing recovered demand with demand no locality chose would break it. It does not. Moreover, in Section 5.3 we show that the same instrument applied to raw housing-unit counts, an outcome the model never touches, reproduces the same density gradient: places with the same predicted inherited demand growth add fewer units where density is higher.

4 Data

Our analysis uses a balanced annual panel of U.S. census tracts spanning 1999 to 2016, covering metropolitan areas nationwide. Census tracts are small, stable geographic units, well suited to studying neighborhood-level variation in housing markets, and the national sample spans the full range of urban densities. Housing quantity is the total stock of dwelling units in a tract, measured by Baum-Snow and Han (2024) from ZTRAX

assessor records of residential parcels; it counts physical units rather than occupied ones, so it is unaffected by vacancy. It also underlies the density measure, defined as units per square mile. The price measure is a tract-level house-price index expressed in real 2002 dollars using the CPI-U.[6] Household income is average household income in the census tract, also deflated to real terms. The tract-level supply elasticities are those of Baum-Snow and Han (2024), converted from decadal cumulative to annual flow terms (Section 3.3). We use them directly, and through the density gradient estimated there we impute an elasticity for every tract they could not measure.

Table 1 reports sample means and standard deviations for select years. Over the sample, the average tract's housing stock grows slowly, by about 0.8 percent per year, while its real house price grows by about 2.8 percent per year.[7]

Sample construction. We exclude tracts with missing key variables. We also exclude county-years affected by documented breaks in the source housing-stock series. The panel's unit counts are built from county records, and in twelve counties the series breaks visibly in a particular year: the recorded stock of most tracts in the county jumps or collapses at once, by around ten to forty percent in a single year, and the series typically freezes at its new level thereafter. County-wide moves of this size in one year are not the secular economic changes we study, and the breaks include upward jumps as well as collapses, so the rule removes spurious booms and spurious busts symmetrically. We drop each affected county from its break year onward, about 2 percent of observations; the largest cases are Cook County (Chicago) from 2014 and Philadelphia County from 2013.[8] All growth comparisons between 1999 and 2016 hold tract composition fixed accordingly.

We set the construction lag to $m = 5$ years.[9] Because of this lag, the first five years of each tract's series are absorbed into the lagged terms and do not enter the regression. The resulting estimation sample contains 299,907 tract-by-year observations.

4.1 The demand elasticity and what the fundamentals imply

The demand elasticity governs how much demand pressure the model attributes to a given observed price increase. We set the elasticity of demand using direct estimates from the literature, and then spend the rest of this section showing that the demand growth it implies is consistent with population, real average

[6] Baum-Snow and Han (2024) build this index from repeat sales of the same physical unit over time, which nets out cross-sectional differences in unit quality by construction; they show it agrees closely with an alternative hedonic index that controls directly for structure characteristics, and use the repeat-sales version as their primary measure for its broader coverage.

[7] Tract composition is held fixed across the endpoints, so the stock figure is equivalently the growth of the total stock across these tracts. The price figure is the growth of the unweighted cross-tract average of the tract price index.

[8] The full list, with break years, is: Tallahassee (Leon) 2002; Anne Arundel 2003; Mobile 2004; Volusia 2004; Sarasota 2009; Ocean, NJ 2009; Montgomery, OH 2012; Knox, TN 2012; Jefferson, AL 2013; Warren, OH 2013; Philadelphia 2013; Cook 2014. Re-estimating the baseline with these county-years left in leaves both results intact, with the interaction at -0.020 and still significant at the 1 percent level; excluding the breaks mainly sharpens magnitudes. All appendix robustness checks are run on the cleaned sample.

[9] The sensitivity of the results to the choice of m is tested directly. Section 8.1 lets the lag rise with density, and Appendix A.9 re-estimates the model under common lags of three, four, and five years. Shorter lags produce larger coefficients throughout that range, so the five-year baseline is, if anything, the conservative choice for the magnitude of our findings.

household income, and the decline in the real mortgage rate over the sample period.

First, U.S. population grew by about 0.86 percent per year, from 279 to 323 million, raising the number of households seeking housing. Second, real average household income in our sample did not rise, so it was neither a source of demand growth nor a meaningful drag on it over this period.[10]

Third, the decline in real mortgage rates lowered the user cost of housing and raised willingness to pay independently of income. Because the user cost depends on the real rate, we net out expected inflation. The nominal thirty-year fixed mortgage rate fell from about 7.4 percent in 1999 to about 3.7 percent in 2016, and ten-year expected inflation fell by about half a percentage point over the same period. As a result the real mortgage rate fell from roughly 5.0 percent to roughly 1.8 percent, a decline of about 3.2 percentage points.

We deliberately stop short of converting that decline into a demand contribution. Doing so would require the demand elasticity, which is the very parameter we are about to set, and the resulting figure would then be unavailable as a check on it. Instead, we run the comparison in the opposite direction: we fix the elasticity from outside our data, and ask what is left of demand growth once population and income are accounted for, and test whether that remainder is the size the mortgage-rate literature says it should be.

Of these three forces, then, only the first two enter as measured demand growth, and together they contribute about 0.86 percent per year, with essentially all of it coming from population. We then set $\varepsilon_d = -0.65$, in the middle of the range estimated from household behavior. Hanushek and Quigley (1980) estimate price elasticities of about -0.45 to -0.64; Polinsky and Ellwood (1979), reconciling micro and grouped estimates, place the elasticity at about -0.7; Mayo (1981), surveying the first generation of demand studies, concludes that most credible estimates fall between -0.5 and -1.0; and Albouy et al. (2016) obtain values of a similar magnitude. Appendices A.5 and A.6 later show this choice is not what drives the result, including when the elasticity is allowed to vary with how substitutable a tract's housing is within its metropolitan area.

With the elasticity fixed, the aggregates become a test the model has to pass rather than an input it was fitted to. The average tract's real house price grew about 2.8 percent per year over the sample and its housing stock about 0.8 percent per year, so the demand curve $\log Q = \log \varphi + \varepsilon_d \log p$ requires demand to have shifted out by about 2.6 percent per year. Population and income account for 0.86 percent of that. The remaining 1.7 percent per year is what the model attributes to the decline in real mortgage rates, and nothing in the construction of that residual used a mortgage rate.

We can therefore ask whether the residual is the right size, in the units the mortgage-rate literature reports. Those studies hold the housing stock fixed and measure how far prices move, so we run our residual through the same experiment. A demand shift of ΔD against a fixed stock raises prices by $\Delta D / |\varepsilon_d|$, which converts 1.7 percent per year into a price response of 2.6 percent per year, or about 56 percent cumulated over the sample.[11] Set against a 3.2 percentage point fall in the real mortgage rate, that is a semi-elasticity

[10] Income in the panel comes from the decennial Census, so it is genuinely observed only twice, in 2000 and in 2010; we therefore measure income growth over that observation window. Deflated by the CPI-U, real average household income was essentially flat, moving from about \$70,066 to about \$69,700, or about -0.05 percent per year.

[11] That price response is a counterfactual: it is what prices would have done had the stock been frozen, which is the experiment the

of real house prices to the mortgage rate of about 11.2 percent per percentage point, within the range of estimates in the literature: Damen and Schildermans (2022) estimate 12.3 and Laufer and Tzur-Ilan (2021) obtain 6 to 10.[12]

This is the demand shift we feed through the model in Section 5.2 to trace price growth across density, where the density gradient in supply absorption makes prices grow faster in dense locations than in less dense ones. We do not tune the supply elasticity to make the typical location reproduce observed average price growth; Section 5.2 reports the average price growth rate the model actually implies and how it compares.

A natural worry about any structural recovery is that the objects it produces are just algebra dressed up as economics. If the recovered demand shifter captures genuine local demand pressure rather than an algebraic residual, it should co-move with fundamentals that never entered the recovery. Appendix A.7 shows that it does: tract income growth and tract population growth over 2000 to 2010 each predict it strongly and positively, with coefficients of roughly 0.8.

5 Results

5.1 Regression results

We apply the specification in (7) to the U.S. tract panel over 1999 to 2016, with demand and supply shifts recovered at the tract level and density defined as dwellings per square mile.[13] The baseline uses a demand elasticity of -0.65, taken from the literature, and the density-declining, time-varying annual supply elasticity of Section 3.3. As explained there, the preferred specification omits tract and year effects and uses robust HAC standard errors. Table 2 reports the estimates.

Supply responds positively to demand, but well short of one-for-one. The coefficient on the demand shift, β_1, is 0.40 and statistically significant: a ten percent increase in demand is associated with roughly a 4.0 percent outward shift in supply in a hypothetical location of negligible density. Because density enters the regression in levels and the interaction below is negative, the response is considerably smaller at a typical tract (Section 5.2). As a result, supply absorbs only a limited share of demand pressure, and the bulk is capitalized into price.

This shift responsiveness is moderated further by density. The interaction between log density and the log of the demand shift, β_2, is negative (and statistically significant) at -0.036, meaning that supply responds less to demand pressures as density rises.[14] Although the per-unit effect looks modest, it compounds across

mortgage-rate literature runs, and it is not a further increase to be added to the 2.8 percent we observe. Because the stock did in fact grow, by 0.8 percent per year, observed price growth came in below what a frozen stock would have delivered.

[12]The 11.2 comes from three steps. First, a house is worth its rent divided by the user cost u, so at a fixed stock a cumulative price rise of 56 percent means u fell by a factor 1.56, that is $\ln(u_0/u_1) = 0.44$. Second, solving $u_0/(u_0 - 3.2) = 1.56$ for the 3.2 point fall gives $u_0 \approx 8.9$ percent, a plausible user cost for 1999. Third, the semi-elasticity is then $1/u_0 \approx 11.2$ percent per percentage point.

[13]Section 8.1 re-estimates the model with population per square mile as the density measure instead; the interaction is, if anything, slightly larger in magnitude.

[14]The interaction also survives replacing the pooled cross-section with metro $\times$ year fixed effects, which compare dense and sparse

the wide range of densities in the data, as Section 5.2 shows.

The standalone effect of density, β_3, is small at 0.005. This estimate indicates that density plays a negligible role independent of demand pressures: absent demand pressure there is little building for density to dampen. Its economically meaningful role operates instead through its interaction with demand.

5.2 Marginal effects and heterogeneity

To illustrate the relationship between supply shifts and density, Figure 2 plots the implied marginal effect of demand shifts on supply, the derivative of $\log \psi$ with respect to $\log \varphi$. This marginal effect is the absorption rate, which is equal to $\beta_1 + \beta_2 \log \text{Den}_{t-m}$, and which we denote ME. The point estimate falls from about 0.261 in the least dense decile to about 0.148 at the median (fifth) decile and about 0.083 in the densest, a reduction of about 68 percent across deciles (Table 3).[15]

The vertical axis has a direct quantitative meaning. A value of 0.26 in the least dense decile means that for every one percent increase in the demand shifter, the supply shifter rises by about 0.26 percent. Supply therefore closes about 26 percent of the demand increase, with the remainder absorbed by price. At the median decile supply closes about 15 percent, and in the densest decile only about 8 percent, leaving the rest to be capitalized into price.

From absorption to price growth. These absorption gaps translate into differences in price growth once fed through the model's market clearing. We hold both the implied aggregate demand shift (Section 4.1) and the supply elasticity fixed across the comparison, and pass them through each decile's supply response alone. Holding the elasticity fixed is deliberate: it removes the separate fact that the supply elasticity also falls with density, so the resulting spread reflects the slower supply shift in dense areas and nothing else. We fix it at the measured sample-median elasticity of 0.036.

The density gradient spreads price growth across deciles, rising from about 2.66 percent in the least dense decile, below the average because it builds the most, to about 3.07 percent at the median and about 3.31 percent in the densest. Relative to the median tract, the densest tract's prices grow about 0.24 percentage points per year faster; relative to the least dense, about 0.65 percentage points faster.

The divergence that follows is driven not by demand or the supply elasticity, which are held common across the comparison, but by how much less the supply of dense locations shifts to meet increases in demand. These wedges are modest per year, but they compound. Cumulated over the seventeen years from 1999 to 2016, the model implies real price increases of about 56.2 percent in the least dense decile, 67.1 percent at the median, and 73.8 percent in the densest. These are the paths the model generates from the absorption gradient alone. They are not the observed decile paths quoted in the introduction, which also

tracts within the same city and year (Appendix A.3). It likewise survives absorbing tract fixed effects outright, roughly twenty-four thousand dummies that wipe out every time-invariant tract-level confounder and leave only each tract's own trajectory: the interaction is -0.027 ($t = -17$).

[15] The confidence bands do not overlap across the extremes, and the gradient remains clearly distinguishable from zero under the more cautious two-way clustered standard errors of Appendix A.2.

reflect differences across deciles in demand growth and supply elasticities, both held fixed here. The gap between the densest and median deciles is therefore about 6.7 percentage points, and between the densest and least dense about 17.7 percentage points; equivalently, the densest decile's cumulative price growth is about 1.31 times that of the least dense.

Separating the shift channel from the slope channel. The wedge above deliberately lets only the shift channel operate, and it is worth being explicit about how the two channels can be separated at all, since both the absorption rate and the supply elasticity fall with density.

In equilibrium, a location's price growth is its unabsorbed demand growth divided by the sum of the absolute values of the supply and demand price elasticities: $\Delta \log p = g_\varphi(1 - \mathrm{ME})/(\varepsilon_s - \varepsilon_d)$, where g_φ denotes the growth rate of the demand shifter. The supply shift channel governs the numerator: how much demand growth is left unabsorbed. The slope channel governs the denominator: how strongly each unit of unabsorbed demand is capitalized into price. This structure makes one point immediately. If supply shifted one-for-one with demand (ME $= 1$), price growth would be zero at *any* elasticity: without a shift gap the slope has nothing to amplify. The slope amplifies the shift failure rather than generating price growth on its own.

The mirror-image limit also holds: a perfectly elastic supply would deliver zero price growth at any shift, since the denominator grows without bound. Sustained price growth therefore requires both a shift that falls short and a curve that is less than perfectly elastic.

The two channels can be separated by counterfactual, exactly as in growth accounting, comparing the densest decile with the least dense. We call the resulting split of the density gap in price growth the shift-elasticity composition. With both differences operating, the densest decile's lower absorption and its steeper supply curve, the model implies that its prices grow about 0.90 percentage points per year faster than the least dense decile's. If the two deciles differed only in absorption and shared a common supply elasticity, the gap would be 0.65 points per year; if they differed only in supply elasticity and shared a common absorption rate, the gap would be 0.27 points, against 0.90 points when both differ. The slope therefore accounts for only about 29 percent of the difference in price growth between the densest and least dense deciles, against about 71 percent for the shift.[16][17]

Table 4 reports the attribution decile by decile, writing each decile's price growth as a common component plus the two channels. The common component, about 2.51 percent per year, is the price growth even

[16] The two channels interact, since the slope amplifies whatever the shift leaves unabsorbed, so a decomposition evaluated at one decile's values leaves that overlap outside the split. We allocate it symmetrically, averaging each channel's contribution over both orderings, which makes the shares sum to one hundred by construction. Evaluated at the least dense decile alone the shares are 68 and 26, with the remainder the overlap, which Table 4 folds into its slope column; the ranking is the same under either ordering. These shares are point estimates at the baseline calibration; Section 8.4 recomputes them under every robustness variant the paper estimates and finds the slope share between 16 and 48 percent across defensible calibrations, with the shift the larger channel under every defensible timing of supply.

[17] For comparison, the observed gap in the raw data is larger, about 1.6 percentage points per year between the same deciles (the annualized counterpart of the cumulative growth reported in the introduction). The counterfactual gap is smaller by construction: it feeds every decile the same demand growth, while in the data demand also grew faster in denser locations. The supply channels alone therefore account for about half of the observed gap, and the shift-elasticity composition here refers to that supply-driven half.

the least dense decile experiences, because supply absorbs only part of demand growth everywhere.[18] The shift and slope columns decompose the extra growth that density adds on top of it. At every density level the shift contribution to that extra growth is roughly twice the slope contribution.

Could the curve instead simply have steepened, without ever shifting out? The attribution above already answers the price side: even in this extreme case, slope differences deliver less than a third of the price gap. But an equally important rejection comes from quantities: along a fixed curve, the least dense decile would have added only about 0.26 percent of stock per year, when it in fact added about 1.5 percent, six times more. Sparse areas built far too much housing for a supply curve that never moved. The densest decile makes the same point from the other side: on a fixed curve it would have added about 0.07 percent per year, when it in fact added about 0.0 percent. Where the cap does not move, the two accounts agree, because there is essentially no shift to explain; where it moves, the fixed curve falls far short of the equilibrium quantities we observe. Building comes mostly from caps being raised, and much less from developers climbing their supply curves.

A complementary exercise asks a different question: not why dense places' prices grow faster than sparse ones, but how much price relief an improvement in either margin, a flatter supply curve or a faster-shifting one, would buy in a given place. Posed this way, the answer depends on the size of the improvement one imagines, since in the limit an infinitely flat supply curve would absorb any demand through the slope alone. We therefore pose it symmetrically: double each margin and compare.

Doubling the least dense decile's supply elasticity would lower its price growth only from about 2.51 to about 2.27 percent per year, whereas doubling its absorption rate would lower it to about 1.61 percent, nearly four times the reduction. In the densest decile, doubling the elasticity would lower price growth from about 3.41 to about 3.34 percent per year, while doubling absorption would lower it to about 3.11 percent, a similar price-reduction ratio.

The quantity margin points the same way. Doubling the least dense decile's elasticity would raise its stock growth from about 0.84 to 1.00 percent a year, while doubling its absorption would raise it to about 1.42 percent; in the densest decile the two would deliver 0.30 and 0.45 percent against a baseline of 0.25. Whether the objective is cheaper housing or more of it, moving the curve buys more than flattening it.

These results suggest that policies that make the supply curve flatter buy far less price relief than policies that let the curve move with demand.

Aggregate stakes. To gauge what the gap in price growth between dense and less dense neighborhoods means in the aggregate, we combine the density gradient with the elasticity of aggregate output with respect to house prices in high-productivity cities estimated by Hsieh and Moretti (2019). That elasticity implies an aggregate output gain on the order of 0.05 to 0.1 percent for each 1 percent reduction in house prices in

[18] It sits slightly below the 2.66 percent of the shift-only exercise because in Table 4 each decile carries its own supply elasticity rather than the common one. The least dense decile's annual-flow elasticity is about 0.077, roughly twice the sample median of 0.036 used in that exercise, and a higher elasticity means a smaller price response to the same unabsorbed demand.

those cities. The natural benchmark for their exercise is the median city, since their counterfactual relaxes constraints in the most constrained metropolitan areas to the median level, so we compare the densest decile to the median. The cumulative price gap of roughly 6.7 percent between the densest and median deciles then implies an aggregate output gain of roughly 0.35 to 0.7 percent of GDP, or $64 to $127 billion per year given U.S. GDP of about $19 trillion over our sample. We offer this as an illustrative order of magnitude rather than a precise estimate.[19]

5.3 The gradient under predicted demand growth

The demand shift used so far is recovered from the same prices and quantities as the supply shift, and part of any location's demand is arguably its own doing: a metropolitan area that permits freely attracts the people who are priced out of one that does not. If accommodated demand is partly a choice, one might worry that the gradient above reflects which places chose growth rather than how supply responds to pressure. This subsection reruns the estimation with a measure of demand that no locality could have chosen.

The measure is built from industry inheritance, using two-digit industries from the County Business Patterns. Before the beginning of our sample, each county hosted a particular mix of industries, and over the years that followed some of those industries grew nationally while others declined. A county heavy in manufacturing in that period was handed weak demand by what happened to manufacturing everywhere; one that started heavy in professional services was handed strong demand. Concretely, we take each county's employment shares in the year just before the sample begins, multiply each share by that industry's national log employment growth to the end of the sample, and sum: $B_c = \sum_s \omega_{s,c}\, g_{s,-c}$, where $\omega_{s,c}$ is industry s's share of county c's employment in that pre-period and $g_{s,-c}$ is the industry's national growth computed leaving county c's own employment out, so that no county's own boom enters its own prediction (Bartik, 1991). The result is the demand growth each area inherited from its starting economic base: fixed before the sample and moved only by national industry fates. Aggregated to the metropolitan level, it correlates 0.41 with the recovered metropolitan demand shift ($t = 5.9$). Inheritance is a real but partial driver of demand, which is what the exercise requires.

We use B in two ways (Table 5). The first is a reduced form: rerun the baseline regression with predicted demand growth in place of the recovered demand shift. Because the prediction is a single growth rate per county rather than metropolitan area, it carries within-metro as well as across-metro variation wherever a metropolitan area spans more than one county; the panel still collapses to a cross-section at the end of the sample. For each tract we regress the supply shift accumulated over 1999–2016 on its county's predicted demand growth, on that prediction multiplied by the tract's log density, and on log density. The sample is the 19,639 tracts in the 344 counties (155 metropolitan areas) matched to the industry data. We cluster

[19] The wedge is measured across tract density deciles while the output elasticity applies to house-price reductions in high-productivity cities, and we do not map the one into the other, so the range is conditional on that correspondence rather than implied by our gradient alone. This is a deliberately conservative wedge: it lets only the shift channel operate, holding the supply elasticity fixed across the comparison, and it spans only the upper half of the density distribution. The same exercise against the least dense decile yields a cumulative gap of about 17.7 percent, implying roughly 0.9 to 1.8 percent of GDP, or about $168 to $336 billion per year.

standard errors by metropolitan area, the more conservative choice given that counties within the same metropolitan area likely share labor-market conditions; Table 5 also reports first-stage strength under county clustering.

The question is the same, whether the supply response to a given push of demand weakens as density rises; only the source of the push has changed, from demand recovered from local prices and quantities to demand inherited from each county's pre-period industry mix. The answer is also unchanged. Supply shifts respond strongly to predicted demand growth ($t = 4.3$), and the response falls with density (interaction -0.13, $t = -4.0$).[20] Read at decile means, a unit of predicted demand moves supply by 0.63 in the least dense decile, 0.22 at the median, and 0.01 in the densest: the same shape as the gradient of Section 5.2 (and in fact a much steeper gradient), falling essentially to zero at the top.[21]

The second use keeps the recovered demand shift as the regressor and uses the prediction as an instrument for it. The instrument remains significant under the more conservative of the two natural clustering choices.[22] The reduced form above asked whether supply responds to inherited demand growth directly. Two-stage least squares instead asks the baseline question with the baseline regressor, but discards every part of the recovered demand shift, and of its density interaction, that inheritance does not predict. Only demand variation no locality chose is left to identify the coefficients. The interaction comes out at -0.096, next to the -0.060 that ordinary least squares gives on the same tract cross-section. Neither figure is comparable to the -0.036 of Table 2: that is the annual panel, whereas both of these are long differences over 1999 to 2016, one observation per tract, with the instrument varying at the county level. What matters is the comparison between them, since removing the part of demand that localities might have chosen moves the estimate away from zero, not toward it, which is the opposite of what a manufactured gradient would do. The gradient does not depend on demand shifts being recovered from equilibrium objects: it is there also when demand is an inheritance.

The same design can also stress-test the model against raw objects that never passed through it. Regressing cumulative housing-unit growth on predicted demand growth, its interaction with density, and density gives an interaction coefficient of -0.19 ($t = -3.7$) with a predicted demand growth coefficient of 1.88 ($t = 4.8$): locations with higher predicted demand growth build more housing, but denser ones build less than sparse ones, given the same predicted demand growth, exactly as the model says. This is a strong test

[20] Four checks pin this reduced form to the variation the design claims. First, reassigning the predicted-demand values at random across the counties of the same metropolitan area, twenty thousand times, an interaction as large as the one estimated arises in only five draws (implying a p value of 0.00025): the result is tied to which county actually inherited which industries, not to anything about metropolitan areas as wholes. Second, computing national industry growth leaving out each county's entire metropolitan area, rather than the county alone, leaves the interaction at -0.14 ($t = -4.0$), so no county's prediction reflects its own boom, nor its neighbors'. Third, dropping the ten largest metropolitan areas entirely leaves it at -0.12 ($t = -3.0$): the gradient is not a story about a few large cities. Fourth, a pre-period placebo: the instrument does not predict tract housing-unit growth over 1990 to 2000 differentially by density (interaction $+0.02$, $t = 0.7$; within-metro permutation $p = 0.90$), nor population growth, and conditioning on the 1990s' own shift-share leaves it null, so whatever the interaction captures in the sample period is not a pre-existing dense-tract trend.

[21] The level of the response is not an absorption share. Predicted employment growth is not identical to housing demand, so a coefficient of 0.63 does not mean that sixty-three percent of demand is absorbed. Only the comparison across densities is informative, since there the units are common and divide out.

[22] First-stage joint F-statistics are 15.3 for the demand shift and 11.5 for its density interaction, clustering by metropolitan area, and 19.3 and 14.8 respectively clustering by county, the level at which the instrument itself varies; in both cases these are at or above the conventional rule-of-thumb threshold for weak identification.

because raw housing-unit growth, predicted demand growth and density all come from the data with no model in the background. The magnitudes agree with the model's own as well. The response to predicted demand growth falls from 1.14 in the least dense decile to 0.26 in the densest, a drop of 0.89, or 78 percent, against the 68 percent decline in absorption that Section 5.2 recovers from the structural estimates.[23]

Like any shift-share design, these estimates are identified off the particular demand differences that national industry reallocation generated across counties over this period, and that is the interpretation to carry: when demand arrives through channels no locality chose, dense tracts absorb far less of it.

6 A theory of elasticities, shifts, and density

The estimates of Section 5 leave a question of interpretation: how does a location's density govern the steepness of its supply curve and how far that curve moves? We now build an environment that answers it, and that delivers an economic explanation for two patterns in the data: a price elasticity that declines with density, and a supply curve that shifts outward more slowly in dense locations. Both of these patterns are functions of a location's buildable cap: the room left under the cap determines the price elasticity of supply, and the rate of movement of the cap is what affects the outward shift. They are not two stories that happen to agree with each other about density; they are one object, seen once as a level and once as a rate of change. Everything below is an elaboration on that sentence. Using the model, we estimate the rate at which the cap grows across tracts and density, and later validate it against independent measures of regulation. The central density relationship, how the supply shift varies with density, is not imposed by the theory but estimated from the data.

6.1 Land, density, and the buildable cap

A location occupies a fixed quantity of land L and contains H housing units, so its density is $\text{Den} = H/L$, the number of units per unit of land, which is the density we measure in the data.[24] It faces a buildable cap $\bar{H}$, the maximum number of units that may be built given current zoning, approvals, and land-use rules, which can be expressed equivalently as a maximum density $\overline{\text{Den}} = \bar{H}/L$ that its land may hold.[25]

What governs how hard it is to add units is not density on its own but how close a location's density

[23]The estimate does not rest on any one industry. Exposure is concentrated, as in most shift-share designs, but dropping manufacturing, which carries the largest share, leaves the interaction at -0.10, and dropping the five largest industries leaves it at -0.17. Moreover, the shift-elasticity composition of Section 5.2 is unchanged in substance when it is rebuilt from these coefficients: the slope carries 13 percent of the density gap in price growth under two-stage least squares, against 18 percent under ordinary least squares on the same cross-section, so instrumenting demand makes the shift channel larger rather than smaller.

[24]The choice of density measure matters little in practice. Merging in tract-level population from the census, log dwellings per square mile and log population per square mile correlate at 0.98 across over 24,000 tracts, so the two are near-substitutes empirically as well as conceptually; Section 8.1 confirms this directly, re-estimating the baseline with population density in place of dwelling density and recovering the same density-dampening result.

[25]Within the dense half of metropolitan areas, the recovered ceiling itself, extracted from quantities and prices, moves with predicted local demand, a Bartik instrument built from each county's pre-sample industry mix and national industry fates, which uses no housing data at all (Section 5.3): caps rose faster where predicted demand was stronger, by 0.28 percentage points a year ($t = 2.8$), and by 0.35 ($t = 3.8$) comparing metropolitan areas at the same density. Whatever the binding limit is, it responds to demand pressure no locality chose, and a physical frontier cannot do that.

is to this maximum, the fill fraction $f = H/\bar{H} = \text{Den}/\overline{\text{Den}}$, which runs between zero and one. A location with f near zero has ample room to build; one with f near one is close to its ceiling. This fraction changes for two reasons: the location builds out, raising Den, or the cap is relaxed, raising $\overline{\text{Den}}$.

6.2 The cost of building and the supply curve

Building is harder the closer a location's density is to its cap, and harder again the more already stands on the ground. The general marginal cost we work with carries both forces,

$$C'(H) = c_0 \frac{H^\theta}{(\bar{H} - H)^\alpha}, \tag{8}$$

where $\alpha > 0$ governs how sharply cost rises as the room to build, $\bar{H} - H$, runs out, and $\theta \geq 0$ governs how much the standing stock itself adds to the cost of the next unit. We develop the model at $\theta = 0$, which integrates in closed form and keeps every expression below readable, and we say plainly what that costs: it assumes the standing stock raises cost only through the room it consumes, not on its own account. Section 6.6 returns to (8), estimates θ, and shows that every result of this paper survives it. Setting $\theta = 0$ and integrating gives the total cost of building H units,

$$C(H) = c_0 \frac{(\bar{H} - H)^{1-\alpha} - \bar{H}^{1-\alpha}}{\alpha - 1}. \tag{9}$$

The scalar c_0 sets the level of construction cost. This cost function follows the literature in having marginal cost rise (Glaeser et al., 2005; Paciorek, 2013; Gyourko and Molloy, 2015). What we add is the cap, which makes the marginal cost rise without bound as building approaches the ceiling. That feature is the essential one.

Other obvious alternatives fail. A cost function without a cap gives a constant elasticity, which the Baum-Snow and Han (2024) gradient rules out. A constant elasticity that differs across locations fits that gradient, but only by making the elasticity a permanent attribute of a place, so that San Francisco and New York City would have been as inelastic when they were small as they are now. And a cap reached at a finite cost predicts that high-demand locations fill it and stay there, which New York City's frozen caps and stable fill fractions contradict (Section 7.3). Appendix A.1 discusses these alternatives in detail.

A cap with an unbounded marginal cost, by contrast, delivers three things the data ask for. First, the elasticity can change with density as the data show it does. Second, the fill fraction, f, stops short of one, as also found in the data. Third, relaxing the cap lowers marginal cost at every amount of construction, shifting the whole curve rather than moving its endpoint. A higher cap raises the profit on each project, which can make mutually beneficial buyouts of incumbents viable where they were not before. This channel is what lets the model match equilibrium prices and quantities without the implausible productivity gains in construction that Louie et al. (2026) require (and reject) of a supply shift (Appendix A.13).

A competitive developer chooses how much to build to maximize the value of construction net of cost. Because construction takes time to complete, the developer responds to the price prevailing when the project is begun, the lagged price p_{t-m}. Facing p_{t-m}, the developer solves $\max_H \{ p_{t-m}H - C(H) \}$, and the first-order condition sets this price equal to marginal cost, $p_{t-m} = C'(H)$. Differentiating (9) gives

$$C'(H) = \frac{c_0}{(\bar{H} - H)^{\alpha}}, \tag{10}$$

which rises as the room left runs out.

Marginal cost depends on the room left, not on the stock already built. A location permitted twelve hundred units and standing at a thousand is scraping its last sites, assembling parcels and contesting approvals; a location permitted five thousand and standing at the same thousand is building on ordinary ones. The cap is therefore not a wall at the end of the path but a price on every step along it, which is why building slows long before the ceiling is reached and why no location ever arrives there. Relax the cap and every quantity becomes cheaper to build, so the entire curve moves outward (down).

Setting $p_{t-m} = C'(H)$ and inverting gives the supply curve, the stock the developer builds at the lagged price,

$$S \equiv H = \bar{H} - \left(\frac{c_0}{p_{t-m}} \right)^{1/\alpha}. \tag{11}$$

Two forces move quantity along and around this curve. A higher price draws the developer further up the marginal-cost curve, raising S: this is how the price elasticity of supply affects the quantity supplied. A relaxation of the cap $\bar{H}$ shifts the whole curve outward at a given price: this is how the supply curve shifts.[26]

6.3 Elasticity and density

The price elasticity of supply measures how far a location moves along its supply curve when the lagged price changes, holding the cap fixed. Partially differentiating (11) with respect to the lagged price at a fixed cap and writing the result in elasticity form gives

$$\varepsilon_s = \frac{1}{\alpha} \frac{1 - f_{t-m}}{f_{t-m}} = \frac{1}{\alpha} \frac{\overline{\text{Den}}_{t-m} - \text{Den}_{t-m}}{\text{Den}_{t-m}}. \tag{12}$$

Here f_{t-m} and Den_{t-m} are the fill fraction and density, respectively, at the lagged date $t - m$, when the building decision is made. Partially differentiating this with respect to density shows that the elasticity falls monotonically as density rises,

$$\frac{\partial \varepsilon_s}{\partial \text{Den}_{t-m}} = -\frac{1}{\alpha} \frac{\overline{\text{Den}}_{t-m}}{\text{Den}_{t-m}^2} < 0. \tag{13}$$

[26] Nothing in Sections 5 and 5.2 uses this cost function. The regressions and the shift-elasticity composition use the Baum-Snow and Han (2024) elasticities and the shifts recovered from prices and quantities given those elasticities, so they hold under any cost function consistent with the measured elasticities, as well as under our Bartik specification. The cost function enters only the recovery of the cap relaxation in Section 6.5, and Section 6.6 shows that even that is robust to writing the standing stock into it.

This is exactly the density gradient in the supply elasticity that the analysis of Baum-Snow and Han (2024) points to.

Equation (13) holds the cap fixed, but along a location's actual growth path the cap can move too, so it is instructive to take the total derivative as well. Writing $\eta \equiv d \log \overline{\text{Den}}_{t-m} / d \log \text{Den}_{t-m}$ for the percent by which the cap is higher in a location one percent denser, and taking the full derivative of Equation (12) gives

$$\frac{d\varepsilon_s}{d \log \text{Den}_{t-m}} = \frac{1}{\alpha} \frac{\overline{\text{Den}}_{t-m}}{\text{Den}_{t-m}} (\eta - 1) . \tag{14}$$

The elasticity falls with density if and only if the cap responds less than one-for-one to densification ($\eta < 1$).[27] Were caps relaxed faster than places build out, room to build would accumulate as density rises and the elasticity would *rise* with density. The declining elasticity gradient that Baum-Snow and Han (2024) estimate is therefore itself equilibrium evidence that caps fail to keep pace with densification.

6.4 Supply shifts and density

This section shows that the regression of Section 5 is what a model of buildable caps predicts, and says what its interaction coefficient means. It also delivers a result worth flagging in advance: the supply shift falls with density even if every locality responds to demand in exactly the same way.

The supply shift is the outward movement of the supply curve at a fixed price, which here is a function of the relaxation of the cap $\bar{H}$. What prompts a cap relaxation, above all, is demand. When demand for a location grows, residents, developers, and local governments come under pressure to permit more building, and the cap is raised in response. Let $g(\Delta D, \text{Den}_{t-m}) = \log(\bar{H}_t / \bar{H}_{t-m})$ denote the proportional cap relaxation over a period, a general function of the demand shift ΔD and of how dense the location is. We leave its form unspecified for now and derive the supply shift for an arbitrary g; we postulate a functional form only afterward, once the structure of the problem is clear, and then show how it relates to the regression of Section 5.

The supply curve (11) sets the built stock at $S = \bar{H} - k$, where $k = (c_0 / p_{t-m})^{1/\alpha}$ is the room left unbuilt at the lagged price; holding that price fixed, k does not change. The base built quantity is the base stock, $S_{t-m} = H_{t-m} = L \, \text{Den}_{t-m}$, and the unbuilt room is $k = \bar{H}_{t-m} - H_{t-m} = L \times (\overline{\text{Den}}_{t-m} - \text{Den}_{t-m})$. Substituting $\bar{H}_t = \bar{H}_{t-m} \, e^g = L \, \overline{\text{Den}}_{t-m} \, e^g$, the land L cancels and the supply shift becomes

$$\log(1 + \Delta S) = \log\left(\frac{\overline{\text{Den}}_{t-m} \, (e^g - 1) + \text{Den}_{t-m}}{\text{Den}_{t-m}} \right) . \tag{15}$$

This holds for any cap-relaxation function g. Because cap relaxations over a period are small, we take a

[27] Our estimates suggest this holds: Section 6.5 puts η at about 0.85, with a bootstrap ninety five percent interval of 0.84 to 0.86.

first-order approximation in g. Expanding $e^g \approx 1 + g$ in (15) gives

$$\log(1 + \Delta S) \approx \frac{\overline{\text{Den}}_{t-m}}{\text{Den}_{t-m}} \times g(\Delta D, \text{Den}_{t-m}). \tag{16}$$

To first order, the supply shift is the cap relaxation scaled by headroom, $\overline{\text{Den}}_{t-m}/\text{Den}_{t-m}$, which is one over the fill fraction f: the shift is thus g/f. Headroom shrinks as density rises, so the same cap relaxation buys a denser place a smaller shift in percentage terms, even though in absolute units it may still add more. That much is arithmetic, and it holds whatever regulators do.[28]

We now postulate a functional form for the cap relaxation, writing it in terms of the demand shift, density, and their interaction,

$$g(\Delta D, \text{Den}_{t-m}) = \beta_1^c \log(1 + \Delta D) + \beta_2^c \log(1 + \Delta D) \log(\text{Den}_{t-m}) + \beta_3^c \log(\text{Den}_{t-m}). \tag{17}$$

Each coefficient has a clear meaning. The pass-through β_1^c measures how strongly a location raises its buildable limit when demand grows. The interaction β_2^c allows this pass-through to depend on density. The hypothesis is that once a great deal has already been built on a location's land, policymakers slow the rate at which they raise the cap. The residents who bear the congestion and disruption of new building are by then numerous and organized, and they oppose it. A negative β_2^c is the formal statement that the demand pass-through is weaker in dense places. The term β_3^c is an autonomous density drift. We impose no values on these coefficients; the theory only names them, and one can hypothesize about the sign of β_2^c. Section 5 does not estimate β_2^c itself; what it estimates is the composite β_2 we derive next.

Substituting (17) into the first-order supply shift (16) multiplies the cap relaxation by the factor $\overline{\text{Den}}_{t-m}/\text{Den}_{t-m}$, which is positive. We measure log density in deviations from its sample mean, so that the average location has $\log \text{Den}_{t-m} = 0$; from here Den_{t-m} is density relative to the sample geometric mean. Expanding this factor to first order about that point gives $\overline{\text{Den}}_{t-m}/\text{Den}_{t-m} \approx c\,[\,1 + (\eta - 1)\log(\text{Den}_{t-m})\,]$, where $c \equiv \overline{\text{Den}}_{t-m}/\text{Den}_{t-m}$ at the average density is a common positive scale and η is the cap-density elasticity of (14).[29] Multiplying out and dropping the second-order terms in $\log(\text{Den}_{t-m})$ gives

$$\log(1 + \Delta S) \approx c\,\beta_1^c \log(1 + \Delta D) + c\left[\beta_2^c + (\eta - 1)\beta_1^c\right] \log(1 + \Delta D) \log(\text{Den}_{t-m}) + c\,\beta_3^c \log(\text{Den}_{t-m}). \tag{18}$$

The cap density has not vanished: it survives as the common scale c on every coefficient. Collapsing the composites into new letters, $\beta_1 \equiv c\,\beta_1^c$, $\beta_2 \equiv c\,[\beta_2^c + (\eta - 1)\beta_1^c]$, and $\beta_3 \equiv c\,\beta_3^c$, the supply shift is a linear

[28] Across locations the inherited cap and the pace of relaxation may vary with density too, and either could in principle reverse the sign. Taking the full derivative of (16) shows that the reversal requires the two cap-side elasticities to sum above one, $\eta + \partial \log g / \partial \log \text{Den}_{t-m} > 1$: caps must rise more than proportionally with density, whether because denser places inherited higher ones or because they relax them faster as they densify. We estimate η at 0.85 and find g falling steeply with density (Section 6.5), so the sum is well below one.

[29] At $\eta = 1$ caps keep pace with density: a place twice as dense was permitted twice as much, and has the same headroom. At $\eta < 1$ they lag behind, so denser places are fuller.

function of the demand shift, density, and their interaction,

$$\log(1+\Delta S) \approx \beta_1 \log(1+\Delta D) + \beta_2 \log(1+\Delta D)\log(\text{Den}_{t-m}) + \beta_3 \log(\text{Den}_{t-m}). \tag{19}$$

The interaction β_2 here is the same coefficient estimated in Section 5, the one reported in Table 2: the theory says what it means, and the regression has already measured it. Two of the three terms in it have signs we already know. The scale c is a ratio of cap density to density, so it is positive by construction, and since the absorption rate is positive at every density decile (Section 5.2), $\beta_1 = c\,\beta_1^c$, the response at the average density, is positive, and so is β_1^c. Only the sign of $\eta - 1$ is open, and whenever $\eta < 1$, which is what Section 6.5 estimates, β_2 is negative even when the cap responds to demand identically everywhere ($\beta_2^c = 0$). This is because the headroom term $c(\eta-1)\beta_1^c$ makes the percentage supply shift fall with density on its own, and any behavioral slowing of the cap response in dense areas ($\beta_2^c < 0$) only deepens it.

Both terms in β_2, headroom and behavior, therefore point the same way, for the reason set out at the start. A dense location stands closer to its cap, which is what $\eta < 1$ implies, so the room the elasticity depends on is smaller; and its cap is raised more slowly, as Section 6.5 estimates, so the movement the shift depends on is slower too. As a result, dense locations have a steeper supply curve *and* a smaller outward shift, so they absorb more of the demand growth in price rather than in quantity.

6.5 Recovering the growth rate of the cap, the g function

An interaction coefficient cannot be checked against anything outside the model. A pace of cap-raising can. This section uses the model and the data to estimate g, the rate at which each location's cap was raised, which is the object Section 7 confronts with survey measures of regulation and with New York's zoning files.

Recovering g also separates two things the regression leaves bundled. A dense tract's supply can shift less for two reasons: mechanically, because the tract is nearly full, so even a generous cap relaxation frees little room to build in percentage terms; or behaviorally, because its cap is being raised more slowly. The regression's interaction mixes the two. The procedure of this subsection separates them, so that the relaxation rates compared across deciles are the behavioral part alone.

Three unknowns stand between the data and g. Two are parameters common to every location: the cost curvature α and the cap response η, how far a location's cap rises as it densifies. The third is specific to each tract: its fill fraction f, the share of its cap already built. The theory supplies three relations, each drawing on a different feature of the data, and we take them one at a time.

The level of the elasticity gives the fill fraction, f. By Equation (12) the supply elasticity is $\varepsilon_s = (1/\alpha)(1-f)/f$, so inverting it gives

$$\overline{\text{Den}}_{t-m}/\text{Den}_{t-m} = 1/f = 1 + \alpha\varepsilon_s. \tag{20}$$

A tract's elasticity is thus a statement about how full it is: an elastic location still has room, an inelastic one has nearly used it up. This is where the tract-level Baum-Snow and Han (2024) estimates enter, and they enter only through their *level*. It is also the only relation containing α, which is why α must be settled before any fill fraction can be read off.

The shape of the gradient gives α and η. One tract's elasticity cannot separate α from f: any α can be made to fit by choosing f to match. What breaks the tie is how the elasticity changes *across* locations of different density. Differentiating Equation (20) with respect to log density, as in Equation (14), gives

$$\frac{d \log \varepsilon_s}{d \log \mathrm{Den}_{t-m}} = (\eta - 1) + \frac{\eta - 1}{\alpha} \frac{1}{\varepsilon_s}. \tag{21}$$

One restriction makes this usable: we assume η is the same everywhere, which restricts the shape of the caps locations have already accumulated and says nothing about the pace g at which any of them is being raised.[30]

With η common, this single equation does the work of two, because we fit it across locations rather than at one of them. Regressing the local slope of the elasticity gradient on the inverse elasticity, the intercept is $\eta - 1$ and the coefficient is $(\eta - 1)/\alpha$: the intercept therefore delivers η, and the ratio of intercept to coefficient delivers α. Fitting across forty density bins gives $\eta \approx 0.85$ (bootstrap ninety five percent interval 0.84 to 0.86) and $\alpha \approx 1.7$ (bootstrap ninety five percent interval 1.53 to 1.94, and 1.5 to 2.2 across bin counts and fit degrees).[31] What identifies these two parameters is the *shape* of the elasticity gradient (how fast the elasticity falls relative to how high it stands), and not its level; the level was already spent on the fill fractions.

The supply shift gives the pace of cap relaxation, g. The last relation is the one that produces g itself. By Equation (16), the supply shift equals the cap relaxation scaled by the cap-to-density ratio, $\log(1 + \Delta S) \approx (\overline{\mathrm{Den}}_{t-m}/\mathrm{Den}_{t-m}) \times g = g/f$, so $g = f \times$ shift. This is the only place the supply shifts recovered in Section 3 enter, and with f already in hand from the first relation, each tract's g follows directly. Nothing in this last step is estimated; it is division. That closes the recovery of f, α, η and g.

The estimated curvature implies fill fractions rising from about 0.45 in the sparsest decile to about 0.76 in the densest, and between 0.70 and 0.78 in the densest across that specification range for α: dense places stand far closer to their caps than sparse ones. It uses nothing but the shape of the Baum-Snow and Han (2024) gradient, and it agrees with independent evidence: Brueckner and Singh (2020) find a fill fraction near 0.9 in the densest U.S. locations, somewhat above the 0.76 we estimate.

Table 6 reports g at the estimated curvature and, in its last three columns, under fill fractions of 0.70,

[30] Letting η rise with density from 0.75 to 0.95 gives $\alpha \approx 1.3$ and a densest-to-least-dense g ratio of 0.11; letting it fall from 0.95 to 0.75 gives $\alpha \approx 2.7$ and a ratio of 0.14, against 0.12 at the constant $\hat{\eta}$ (Table 6). However the inherited ceilings scale with density, dense caps are raised at about a tenth the pace of sparse ones.

[31] We smooth the gradient with a quadratic in log density before taking its local slope; the estimates are stable across bin counts and smoothing choices. Equation (21) is exact rather than an approximation, and it carries an overidentifying implication we can check: the intercept and the coefficient must share a sign, which the estimates do.

0.80 and 0.90 imposed on the densest decile instead. The level of g moves with the fill fraction, since a fuller location has less headroom and thus needs a larger cap relaxation to produce the same shift we estimate from the data. However, the ratio of g in the densest decile to g in the least dense varies very little.

Caps in the sparsest deciles were relaxed at about 0.1 percent to 0.2 percent per year over the sample; in the densest decile, at about 0.01 percent per year. The cap in dense places does not merely sit closer to what is already built; it is being loosened at roughly one-tenth the pace, which is the paper's central finding restated in the theory's own units.

What the recovered pace does not by itself settle is why the cap in dense places moves so slowly: whether remaining capacity is physically exhausted, so that no permission could add much, or whether the machinery that grants permission withholds it. Prices and quantities cannot apportion the two, and we do not claim to. They can, however, bear on the purely physical reading: within the dense half of metropolitan areas, the recovered ceiling rises faster exactly where demand pressure, predicted from pre-sample industry inheritance, is stronger (footnote 25), and a fixed physical frontier cannot respond to demand pressure no locality chose.

The survey evidence of Section 7 shows that regulation plays a systematic part: measured regulation loosened least where the recovered caps rise most slowly. The legal cap itself is measurable at the parcel level in one city, New York City, and Section 7.2 checks the recovered pace against it directly.

6.6 A cost function in which the standing stock matters

Section 6.2 introduced the general marginal cost (8) and then set $\theta = 0$. This subsection returns to that choice: it writes the model with θ free, shows the parameter can be estimated, estimates it, and reports what changes and what does not.

The limitation with $\theta = 0$. Marginal cost in (10) depends on the room left and not at all on what already stands. A location permitted three thousand units and standing at a thousand, and one permitted twelve thousand and standing at ten thousand, both have two thousand of room, and both are therefore assigned the same marginal cost. However, if all parcels are already built on, then netting an extra unit in the second means clearing ten times as much occupied building, and those buyouts are precisely what Appendix A.1 argues makes the last increments costlier.

Restoring the exponent ($\theta > 0$). This is the force (8) already carried and that we set aside by setting $\theta = 0$. Writing it back in, marginal cost is multiplied by a power of the standing stock,

$$C'(H) = \frac{c_0\, H^{\theta}}{(\bar{H} - H)^{\alpha}}, \qquad C(H) = c_0\, \bar{H}^{\,1+\theta-\alpha} \int_0^{H/\bar{H}} x^{\theta}(1-x)^{-\alpha}\, dx, \tag{22}$$

with $\theta \geq 0$, and $\theta = 0$ returning (9). The integral is an incomplete beta function, elementary whenever θ is a whole number; at $\theta = 1$ it is

$$C(H) = c_0 \left[\bar{H} \, \frac{(\bar{H} - H)^{1-\alpha} - \bar{H}^{1-\alpha}}{\alpha - 1} - \frac{\bar{H}^{2-\alpha} - (\bar{H} - H)^{2-\alpha}}{2 - \alpha} \right], \tag{23}$$

whose first term is (9) scaled by the cap. Under this form the two locations of the example above would no longer share a marginal cost. With the same room left, the one standing at ten thousand would face a marginal cost 10^{θ} times the one standing at a thousand.

How we estimate θ. A factor in the stock leaves a trace in the supply elasticity, the fill fraction and the cap relaxation:

$$\varepsilon_s = \frac{1 - f}{\theta(1 - f) + \alpha f}, \qquad f = \frac{1 - \theta\varepsilon_s}{1 + (\alpha - \theta)\varepsilon_s}, \qquad g = \frac{\Delta S}{1 + (\alpha - \theta)\varepsilon_s}, \tag{24}$$

which are the expressions of Section 6.2 with α replaced by $\alpha - \theta$ wherever a fill fraction or a cap relaxation is read off an elasticity.

The three relations of Section 6.5 still do their work, with (24) in place of their $\theta = 0$ counterparts. The extra parameter needs one more moment, so estimation rests on two doing different jobs. The first is the national elasticity gradient of Section 6.5. Differentiating the first expression in (24) with respect to log density gives its local log-log slope, $(\eta - 1)[(1 - \theta\varepsilon_s)/(\alpha\varepsilon_s)][1 + (\alpha - \theta)\varepsilon_s]$, which collapses to (21) at $\theta = 0$. However, this only traces out a curve of (α, θ) pairs, one curvature for every candidate θ, and never selects a point on it.

The second moment supplies what the gradient cannot, which is a level. No cap is on the record outside New York City, so no fill fraction is available outside of it, and an elasticity on its own cannot say whether a location is nearly full with costs that climb gently or half empty with costs that climb steeply. The gradient leaves α and θ tied together, so one more fact is needed to pin them down, and New York City's 2016 zoning records supply it.

Over the 77 tracts we can match, measured fill averages 0.65. At $\theta = 0$ the model says those tracts should be 0.81 full on average, which is too full. Raising θ lowers the fill the model predicts, and there is exactly one value at which prediction meets measurement. The target is the average fill across those tracts, not each tract's fill separately:[32] $\theta = 0.37$ with $\alpha = 3.7$, a bootstrap ninety five percent interval for θ running from 0.26 to 0.47. The 2011 vintage of the same records gives a similar value for θ of 0.30.

What it changes. Table 7 sets the $\theta = 0$ baseline and the analysis with the new cost function side by side. Panel B answers the question a reader will be asking, which is how the results change once $\theta > 0$. Nothing in Sections 5 or 5.2 moves at all, the two columns being exactly identical, because the regressions and the

[32] Tract by tract the fit is positive: at the calibrated values the model's fill correlates 0.41 with the measured fill across the 77 tracts. The model also compresses the spread, with a standard deviation of 0.10 against 0.24 in the records, partly because measured fill exceeds one in tracts holding buildings larger than 2016 zoning allowed. The calibration targets the mean, and the cross-section is a diagnostic rather than a moment we fit.

shift-elasticity composition never use the cost function. They use the Baum-Snow and Han (2024) elasticities and the shifts recovered from prices and quantities given those elasticities.

Panel C is where the extension has some impact. Fill fractions fall, since the same elasticity now implies locations stand further away from their caps, and the recovered cap relaxations, g, fall with them. The sparsest decile's g falls from 0.112 percent a year to 0.072 percent and the densest decile's from 0.014 percent to 0.011 percent, both of the same order of magnitude as the baseline with $\theta = 0$. The densest decile's cap moves at 0.12 of the sparsest decile's rate under our function and at 0.15 under the extension, which is very similar.[33]

Why this is not our baseline specification. There are two reasons. First, it cannot be estimated without New York City. The elasticity-density gradient alone fixes only a curve of (α, θ) pairs, so the level of fill comes from one city's zoning records, and the paper's headline specification should not rest on one city. Second, it would complicate the intuition and the explanations without adding anything substantive to the analysis: Panel B is identical by construction and Panel C moves the cap-relaxation ratio marginally from 0.12 to 0.15. A reader who prefers the extension can read Table 7's second column throughout and reach every conclusion the paper reaches.

7 Cross-validation with outside measures of regulation

Section 6.5 estimated the rate at which each location's buildable cap was raised, g, and found that it falls sharply with density, from about 0.1 percent to 0.2 percent a year in the least dense deciles to about 0.01 percent in the densest, roughly a tenfold difference. Every input to that estimate came from inside the model: prices, quantities and the elasticity gradient, with no information about rules, zoning, or regulations. This section asks whether regulation, measured directly and from outside the model, moved the way that estimate implies.

We use two measures from outside the model, and they answer different halves of the question. The Wharton survey of local land-use regulation covers metropolitan areas nationwide, but records the machinery of regulation rather than the cap itself; New York City's zoning records give the cap itself, lot by lot, but for that one city only. We take the survey first and turn to the zoning records in Section 7.2.

The first of the two is the Wharton survey of local land-use regulation (Gyourko et al., 2008), the standard direct measure of regulatory stringency. It was fielded twice, in 2006 and in 2018, and 784 communities answered both waves, 360 of them in our estimation metropolitan areas. Following the same communities across the two waves shows how the regulatory process changed over the period in which we estimate the supply shifts. We cannot use the headline index for this, because it is standardized within each wave and so cannot register absolute change. We use instead the components recorded in identical units in both

[33] The curvature, α, is not the same parameter in the two columns and there is no reason to expect it to be. Under $\theta = 0$, α alone carries the whole climb into the cap; under $\theta > 0$ the climb is shared, and what converts an elasticity into a fill fraction or a cap relaxation is the difference $\alpha - \theta$, which is 1.66 in the first column and 3.4 in the second.

questionnaires: review times and an approval-delay index in months, counts of the bodies whose approval a project needs, and shares of communities imposing a given rule.

Every one of these is something people decide. How many months a review takes, how many bodies must sign off on a rezoning, whether a community caps the permits it issues in a year, whether it requires lots of at least an acre, whether it attaches an affordable-housing condition: none is a fact about the land or a limit of construction technology. They are choices, made and remade by councils and boards, and reversible by the same route. That is what makes the recovered cap a plausible object rather than a residual. The shift we attribute to it asks nothing to become physically easier to build; it asks only that the people who decide what may be built decide differently. That distinction is what Appendix A.13 turns on, where the supply shift Louie et al. (2026) reject as implausible growth in construction productivity can be rationalized as ordinary movement of the buildable cap.

Throughout we sign every measure so that a *positive* number means regulation got *looser*, the same direction as a positive g. A positive entry for approval delay therefore means months of delay were removed, and a positive entry for a share means fewer communities impose the rule. Where a single number is wanted we build a composite the way the Wharton index itself is built. We take the seven components observed for every matched metropolitan area, standardize each across metropolitan areas, and weight them by the loadings of the first principal component rather than equally. The resulting index is itself standardized, so it is measured in standard deviations.

What it measures, though, is movement over time. Each component enters as the difference between the two waves in a quantity recorded in the same units in both, so the underlying object is a change. That is what separates it from the headline Wharton index, which is standardized within each wave and can therefore rank metropolitan areas against each other at a moment but cannot say whether regulation loosened; here the standardization only fixes the scale of a comparison that is already between periods.

Splitting by density. Start with the comparison the recovery most directly implies. If caps are raised more slowly where density is high, regulation should have loosened less there. Table 8 splits the matched communities at the median density and compares the two halves, and regulation did loosen less in the dense half. Overall approval delay fell by about a month in the sparse half while staying flat in the dense half. Minimum lot sizes above one acre, a direct cap on density, spread three times as fast in the dense half, ten percentage points against three. And the number of bodies whose approval a rezoning requires rose everywhere, but significantly more in denser metropolitan areas ($t = 2.6$). The final column states each of these as a single loosening gap, and it is negative for every measure.[34]

Eight gaps sharing a sign could still be luck, and the odds cannot be read off row by row because the components are correlated: the three timing measures move together, so eight rows carry fewer than eight independent draws. Rather than model that correlation, we build a test that carries it automatically. We

[34]This is a comparison between communities in the sparse and dense halves of the sample, so it speaks to the neighborhood-level gradient of Section 5 rather than to any comparison of metropolitan averages.

take the ninety-four metropolitan areas behind Table 8 and reassign them to the dense and sparse halves at random, ten thousand times. Each metropolitan area moves as a whole: its communities keep their actual answers, so their eight component changes travel together in every draw, exactly as correlated as they are in the data; only the split is random. All eight gaps come out negative in about 1.6 percent of these relabelings (Table 8): the alignment is something a random split of the same correlated data produces only rarely. As a check that the correlation is indeed being carried, collapsing the three timing measures into their average and re-running the test on the resulting six components gives $p = 0.021$, essentially the same answer.

Splitting by density and demand. That comparison splits on density but ignores demand, and the theory says demand is what makes density bite. A location nobody wants to build in has no reason to move its cap however built out it is; it is only when demand arrives that being close to the cap constrains anything. The sharper test is therefore two-dimensional. We cut metropolitan areas at the median on density and, separately, at the median on the recovered demand shift, giving four groups (Table 9). The strong-demand half grew at about 3.4 percent a year and the weak half at about 1.0 percent, so both halves are growing; weak means slow, not shrinking.[35]

Both cuts use objects recovered for the metropolitan area itself: we aggregate its tracts to metropolitan totals of dwellings and land and to dwelling-weighted mean prices and elasticities, then apply the recovery of Section 3 to those aggregates. This is not the same as averaging tract-level shifts, because the recovery is nonlinear in prices and quantities.

Table 9 reports the comparison the theory makes: within each demand half, how far the dense group's loosening fell short of the sparse group's. Among metropolitan areas facing strong demand the dense group loosened half a standard deviation less than the sparse group, a gap significant at the 5 percent level. Among those facing weak demand the gap is 0.17 standard deviations, a third the size and indistinguishable from zero. That is the pattern density-dampened absorption implies: being built out matters when there is something to absorb, and not otherwise.

Two further checks guard this comparison, and both are rows of Table 9. First, with forty-five metropolitan areas in each half, standard errors lean on approximations, so we also test the gap without any formula. Within the strong-demand half, eighteen of the forty-five metropolitan areas are dense, in that they sit above the median density, and twenty-seven are sparse. We discard that dense-sparse assignment and draw a fake one at random, re-labeling eighteen of the forty-five as dense regardless of their actual density, and recompute, each time anew, the dense-minus-sparse loosening gap that the fake assignment implies. Twenty thousand such draws trace out the gaps that arise when density genuinely does not matter, and a gap as large as the observed half a standard deviation arises in only 2.6 percent of them. The same exercise in the weak half confirms that its small gap is as unremarkable as it looks: 39 percent of random assignments

[35] The two sides of the comparison span different windows. Regulation is compared between the 2006 and the 2018 survey; demand is the shift our panel delivers over 1999 to 2016, and density is taken at 2006, so that the split is fixed before the change being measured. We do not restrict demand to the years after the first wave, because regulation is slow. An ordinance rewritten in 2012 answers pressure that had been accumulating for a while, not the demand of that year.

produce a larger one.

Second, the median split is a choice, so we drop it and let density enter as a continuous variable. Within the strong-demand half, loosening declines in log density itself, at -0.38 standard deviations per log point ($t = -3.05$); within the weak half the slope is nil (-0.02). The interaction the theory predicts is in the data without any cut, and only where demand supplies the pressure.

7.1 The comparison under predicted demand

The strong- and weak-demand halves above were defined by the recovered demand shift, which invites the same question as in Section 5.3: if accommodated demand is partly a locality's own choice, the strong half may be selected on choices rather than on pressure. We therefore redo the split with inherited demand. We re-rank the metropolitan areas that match the industry data, this time by the demand growth predicted from each metropolitan area's constituent counties' industry mix, call the top half strong, and change nothing else: within each half we again compare how much regulation loosened in the dense group against the sparse group.

The lower panel of Table 9 reports the result. Among metropolitan areas whose inherited demand was strong, the dense group loosened 0.45 standard deviations less than the sparse group, and a gap that large arises in only 2.0 percent of twenty thousand random assignments of the dense label. Among those handed weak demand the gap is 0.02 and unremarkable ($p = 0.95$). Dropping the split and letting density enter continuously gives a slope of -0.20 standard deviations per log point of density in the strong half ($t = -2.3$) and -0.02 in the weak. Regulation eased least where demand pressed on density, and it did so also when the demand doing the pressing was fixed by the industries a place hosted in 1998, the year before the sample begins, before any of the rule changes the surveys record.

Overall, the data suggest that measured regulation, followed through time, moved the way our recovered cap growth says it did: the machinery that changes what may be built eased more where demand grew and density was low than where demand grew and density was high.

7.2 Measuring the cap directly: New York City's zoning records

The survey evidence is still one step removed from the object itself: the Wharton indices measure the machinery of regulation, not the quantity it caps. For one city, the cap itself is on the record. New York City's PLUTO files list, for every tax lot, the maximum floor area the Zoning Resolution permits; when a neighborhood is rezoned the number moves, and when zoning stands still it does not move by even a single square foot. Summing lot area times allowed floor-area ratio over each tract's residentially zoned lots, in archived vintages of the file, gives the direct administrative counterpart of the buildable cap we recover from prices and quantities, for a city that sits squarely in our densest national deciles.[36]

[36] We use three vintages of the file: 2002, the earliest release in the city's archive, 2011, and 2016, the sample's final year. The 2002 and 2011 files carry a single maximum-allowable-FAR field; later files split it into residential, commercial, and facility FARs. The residential

The record says what the recovery says. From 2002 to 2016, essentially our sample window, the median tract's allowed residential capacity grew by 0.000 percent a year, exactly nothing. The citywide total moved from 5.28 to 5.24 billion square feet across 1,799 tracts, through the era of the city's celebrated rezoning program, whose upzonings and downzonings very nearly net out. Between 2011 and 2016 alone, the capacity of 25 percent of tracts did not change by even a single square foot. Our recovered cap growth for the densest deciles is a few hundredths of a percent a year; the measured caps of the densest American city grew, if anything, more slowly still.

The measured cap also moves the way the recovered one does: within the city, capacity grew faster in the tracts whose recovered demand pressure was stronger ($t = 1.8$, holding density fixed), the same demand response the recovered ceiling shows in Section 6.[37] New York City is just one city, and we do not offer it as a representative one: it is the extreme of density, where the recovered cap may be most contested, and the only city whose zoned capacity is recorded lot by lot far enough back to check.

Moreover, the frozen record suggests that it is law rather than geology that drives the caps. In 2016, 25 percent of the city's residentially zoned lots carried buildings standing above the floor area 2016 zoning allowed on their lot, structures that could not legally be rebuilt at their own size if demolished. This is the residue of downzoning: the 1961 Zoning Resolution and the contextual rezonings that followed set allowed densities below what had already been built across much of the city, so in many frozen tracts the cap does not merely fail to rise; it sits below the built stock itself. Whatever stands there was once possible to build; what froze between 2002 and 2016 was permission, not possibility.

7.3 Do the data agree with the estimated fill fractions?

The fill fraction f is one of the model's central unobservables: every tract's fill fraction is inferred from its supply elasticity rather than seen. The same New York City records from Section 7.2 let us look at it directly, because they report both halves of it, the residential floor area standing on each lot and the floor area its zoning allows. We use New York City alone because no other city publishes it in a way that can be used to compute f.[38]

Dividing the residential floor area standing in each tract by the floor area its zoning allows gives f directly, with no elasticity and no model in between. Across the city's 2,052 residentially zoned tracts, the

field is the exact continuation of the old one: on the 723,251 residentially zoned lots whose zoning district is unchanged between the 2011 and 2017 files, the two coincide to within 0.01 in 100.0 percent of cases. The measure is legal floor area on residentially zoned lots, so it abstracts from height limits that occasionally bind before floor area does and from capacity transferred between adjacent lots, which nets out within a tract. The city's largest capacity additions of the period came from rezoning manufacturing and commercial land for housing, and these are counted: once rezoned, those lots are residentially zoned and enter the measure. The only housing capacity not counted is what commercial zoning itself allows, so if anything the measure understates the city's capacity and its growth.

[37] The predicted-demand version of this comparison has no power within a single city. The instrument of Section 5.3 varies at the county level; New York City has five counties, whose inherited industry mixes are nearly identical (B from 0.17 to 0.24), and an exact permutation test over all 120 reassignments of the borough values leaves the response indistinguishable from chance ($p = 0.21$). The cross-metropolitan version of the experiment, identified off hundreds of counties, is the one reported in Section 6.

[38] The alternatives fail for different reasons. The National Zoning and Land Use Database (Mleczko and Desmond, 2023) reports one banded figure per municipality, with an open top category, so it does not vary across tracts. The state files of the National Zoning Atlas vary by district but cover few states and record numeric caps for few dense districts, where form-based codes state no limit in units per acre. Song (2025) recovers minimum lot sizes nationwide, but per lot rather than per tract and for single-family housing only.

mean tract's fill fraction stood at about 0.71 in 2016 and at approximately the same level in 2011. In the densest and most expensive city in the U.S., with its caps frozen through the whole window, the average neighborhood used only that share of what it was permitted, and stopped.

That number is hard to explain if the ceiling were reached at an ordinary marginal cost. Let demand grow long enough against a fixed cap and a location with bounded marginal cost fills to the cap and stays there; if the wall sits at the level of the individual lot rather than the tract, every lot fills in turn and the tract still converges. New York City has had both the demand and the frozen caps. Our own estimates put its demand growth at about 3.6 percent a year between 2011 and 2016, or about 1.1 percent a year from 2004 to 2016 if one allows building to lag demand. Yet its fill fractions did not converge on the cap.[39]

The estimated fill fraction levels also line up with what the elasticities imply. For the 77 New York City tracts our estimation sample covers in 2016, the fill inferred from the supply elasticity averages 0.81 against a measured 0.65 (and 0.81 against 0.68 in 2011),[40] and the model's densest national decile sits at 0.76. The inference therefore lands in the right region. In fact, the fill fractions we use imply a larger cap relaxation within the model than the measured ones would, since $g = f\,\Delta S$.

8 Robustness

Our two central results are estimated on objects the model recovers rather than observes, so we now ask how much they depend on the assumptions behind that recovery. Not everything does. Section 5.3 re-runs the estimation on demand predicted from pre-sample industry mix, which uses no housing data at all, and applies that same instrument to raw housing-unit counts, an outcome the model never touches; Section 7 confronts the recovered caps with survey and zoning records measured entirely outside the model. What is left to test is the supply-shift recovery itself, which depends on the supply elasticity.

The subsections that follow do four different jobs. Section 8.1 varies the model's inputs and the regression's specification one at a time, which tests the magnitude of the estimates. Section 8.2 stress-tests the input with the most potential to reverse their sign, the way the supply elasticity varies with density. Section 8.3 leaves the data behind and runs the whole procedure on artificial markets whose supply behavior we set ourselves, which tests the machinery rather than the assumptions. Section 8.4 asks whether the headline shift-elasticity composition, and not only the interaction coefficient, survives the same alternatives. Table 10 summarizes the checks of this section.

The appendix reports a second, complementary set of checks, in the three groups its introduction sets out. The first concerns inference: standard-error clustering (Appendix A.2) and metro and metro $\times$ year fixed effects (Appendix A.3). The second concerns the demand side: demand spilling across tract boundaries (Appendix A.4), the demand elasticity itself, varied by assumption (Appendix A.5) and with an independent

[39] The window is five years because the 2002 file's built floor area is not comparable with the later vintages.

[40] That measured figure is for these 77 tracts alone, the ones our estimation panel covers, while the 0.71 quoted above is the mean across all 2,052 residentially zoned tracts in the city.

measure of how substitutable each tract's housing is for its neighbors' (Appendix A.6), and a validation of the recovered local demand shifter against real income and population growth (Appendix A.7). The third concerns the other inputs to the recovery: the decadal-to-annual elasticity conversion (Appendix A.8), a closer look at the construction-lag and boom-bust checks below (Appendices A.9 and A.10), imprecision in individual tract elasticities (Appendix A.11), and mismeasurement of the observed prices and quantities themselves (Appendix A.12).

8.1 Specification choices

Table 10 collects the checks that vary the model's inputs and the regression's specification one row at a time. None of them disturbs the density interaction.

Row two adds year effects, which absorb the common, system-wide component of the demand shift and leave the tract-specific deviation; the interaction is essentially unchanged and the demand-shift coefficient barely moves. That check does double duty, since it also answers whether the interaction is an artifact of pooling the 2000s boom with the post-2007 bust: year effects absorb the cycle entirely, and Appendix A.10 re-estimates separately on the 2004–2007 boom and the 2008–2016 bust and recovery, finding the interaction negative and significant in both.

Row three holds each tract's supply elasticity fixed at its year-2000 value, removing the within-tract steepening the baseline allows; the interaction remains negative and highly significant at -0.027, so the dampening is identified primarily across tracts of different density rather than within a tract over time.

Row four lets the construction lag rise with density, from three years in the least dense decile to nine in the densest, since vertical construction takes longer to complete, as the completion and approval timelines in Appendix A.9 document; the interaction comes out larger in magnitude than the baseline's, not smaller.

Row five replaces predetermined density with contemporaneous density, which is mechanically correlated with the dependent variable because both are built from the same period-t stock; the interaction rises by about a third, confirming the baseline as the conservative choice.

Finally, we change the definition of density. Our measure counts dwellings per square mile, because the theory's object is how built-out a location's land is. Re-defining density as census population per square mile instead gives an interaction of -0.039, and fixing population at its predetermined 2000 level gives -0.047, both larger in magnitude than the baseline -0.036 and at least as precisely estimated. The dampening is about crowded places, however the crowding is counted.

8.2 Mismeasurement of the elasticity gradient

Section 3.5 identified the input with the most direct potential to overturn the sign of the interaction: the relationship between the supply elasticity and density. The gradient we use is estimated by Baum-Snow and Han (2024), but any estimate carries error, so we stress-test it directly. We rescale the slope of the fitted

gradient (6), pivoting on the mean log density so that the average elasticity is unchanged, and re-run the entire recovery and regression at each scaling. The scalings span far beyond any plausible estimation error: a slope of zero assumes the gradient is entirely spurious, giving every tract the same elasticity, while two and three times the estimated slope assume it is badly understated.

The last three rows of Table 10 show the interaction is essentially insensitive: β_2 is -0.031 with no gradient at all, -0.036 at twice the estimated slope, -0.029 at three times, and highly significant throughout. The density dampening is identified by the co-movement of prices, quantities, and density in the data, not by the elasticity schedule we assume. No error in the gradient, within or well beyond its plausible range, generates the result.

That stress test tilts the whole schedule at once. The complementary worry is idiosyncratic error in individual tracts, since Baum-Snow and Han (2024) publish point estimates but no tract-level confidence intervals. Perturbing every tract's elasticity by an independent draw from a deliberately extreme ± 0.5 band and redoing the entire recovery a hundred thousand times leaves the interaction negative in every replication, with a 95 percent range of $[-0.037, -0.035]$ around the baseline -0.036 (Appendix A.11). Independent errors across twenty-four thousand tracts largely average out; the dangerous form of elasticity error is the systematic one, and that is what the stress test before this one already ruled out.

8.3 A placebo Monte Carlo

The checks so far change one input at a time and show that the result survives. However, a more basic worry is still on the table. Both recovered shifters are built from the same observed prices and quantities, the supply shifter depends on assumed elasticities, and the regression then relates one constructed object to another. Could the negative interaction be a by-product of that construction rather than a fact about housing markets? Nothing we do to the real data can answer this, because in the real data nobody knows the true answer. In an artificial market we do know it, because we choose it. So we build markets whose supply behavior we have set ourselves, run the whole procedure on them, and compare what it reports with what we put in.

Each artificial market is a copy of our actual sample. Every tract keeps its real initial stock, initial price, density and supply elasticity; only the shocks that hit it are invented. We give each tract a path of demand growth drawn to resemble the demand paths we recovered from the real data, and we then set, by hand, how strongly its supply responds to that demand growth. In every design below each tract gets the *same* response coefficient as every other tract, however dense it is, so the true interaction is zero because we made it zero. Prices and quantities follow from the model's equilibrium. We apply the recovery formulas of Section 3 to those artificial series exactly as we apply them to the real ones and re-estimate the interaction regression, 300 times per design.

One point needs settling before running any simulations, since it determines which designs are worth running. If the simulated world uses the same elasticities the recovery uses, the recovery is the exact algebraic

inverse of the process that generated the data. When $\hat{\varphi}$ and $\hat{\psi}$ are written out against the equilibrium that produced p and Q, the price terms cancel identically, and in our simulations the recovered shifters equal the true ones to machine precision. A correctly specified design therefore *must* return $\beta_2 = 0$; it verifies that the code implements the algebra, and it is not evidence about housing markets. That is design A in Table 11, reported as the reference point and nothing more.

What is worth running is a world the recovery gets *wrong*, since there the answer is not known in advance. Two such worlds matter, and they are the two ways the pipeline could fabricate a density gradient out of nothing.

An elasticity gradient that is not really there. Design B gives every simulated tract the same supply elasticity while the recovery goes on imposing the estimated gradient. Now the recovery no longer inverts the world: the recovered supply shift picks up an extra term, the gap between the assumed and the true elasticity multiplied by past price growth, and that gap is systematically signed by density because the assumed gradient falls with density. A density-correlated term therefore enters the dependent variable by mistake, which is exactly the channel Section 3.5 flags as the most direct threat to our sign. It does not: the interaction comes out $+0.007$, the *opposite* sign from ours. Imposing a gradient where none exists pushes the estimate the other way, so it could not have produced our result.

Error in the data that is worse where we find the effect. Design C is an especially hostile one. The recorded housing stock enters the demand shifter and the supply shifter through the same term, so an error in a tract's records moves both together; regressor and dependent variable then share an error, and least squares is no longer unbiased. If that error is larger in sparse tracts than in dense ones, the resulting bias varies with density and looks like our finding. Nothing cancels here, and the size of the damage is a genuinely open question. We impose the pattern deliberately, with recording noise falling from 5 percent of the stock in the sparsest tracts to 1 percent in the densest. It estimates the interaction coefficient at -0.002, about six percent of ours. Measurement error built deliberately to mimic our finding therefore cannot manufacture that finding where none exists.

8.4 How stable is the shift-elasticity composition?

The checks above concern the interaction coefficient. A separate question is how the headline shift-elasticity composition, which says that about 71 percent of the density gap in price growth is driven by the shift and the rest by the slope (Section 5.2), moves under the same alternatives.

We recompute it under the four inputs that move it: the demand elasticity (Appendix A.5), the decadal-to-annual conversion (Appendix A.8), the elasticity gradient, and the construction lag (Appendix A.9). The demand elasticity leaves the slope share near its baseline 29 percent at $\varepsilon_d = -1.0$, where it is 30, and raises it to 46 at $\varepsilon_d = -0.3$. The shares move slightly more aggressively with the elasticity conversion, which turns

the decadal stock elasticity into an annual one, from 23 percent at a divisor of 15 to 44 at no conversion at all, and with the gradient, 19 percent at half the estimated slope, 41 at double, and 48 at triple.

The one choice the shares genuinely lean on is the timing of supply, so there we do not assume a number; we take the schedule from the completion timelines of Appendix A.9, which lengthen with density, from about three years for the low-rise construction typical of sparse tracts to about nine for dense high-rise projects. Assigning each tract the lag its density implies, three years in the sparsest decile rising to nine in the densest, delivers a slope share of 16 percent, more shift-heavy than the baseline, not less. A uniform three-year lag applied to every tract gives the same 16 percent, so the schedule's shift-heavy answer is coming from its short lags rather than from the variation itself; the baseline uniform lag of five years gives 29 percent.

The point estimate is 71 to 29; the robust statement is that the shift carries the majority of the gap under every defensible calibration, and the introduction's policy comparison is never overturned: doubling the densest decile's supply elasticity always lowers its price growth by less than doubling its absorption rate does, at most 0.85 times as much across the calibrations we examine.

8.5 Summary: the findings survive

Both central findings, that supply shifts by less than demand and that it shifts by less still where density is high, survive the checks we subject them to in this section and in the appendix. They survive the inclusion of year effects, a supply elasticity held constant over time, rescaling the elasticity-density gradient from zero to three times its estimated slope, perturbing every tract's elasticity independently within a ± 0.5 band across a hundred thousand replications, and a construction lag that rises with density. They survive tests of the machinery itself (Section 8.3) and mismeasurement of the observed prices and quantities. They likewise survive more cautious standard errors, comparisons made only between neighborhoods of the same city, and the split into the years before and after the 2008 crash (Appendices A.2, A.3 and A.10), alternative and income-dependent demand elasticities, and a demand elasticity that varies with how substitutable a tract's housing is (Appendices A.5 and A.6).

The headline shift-elasticity composition is less stable, moving between 16 and 48 percent across defensible calibrations, but what it is there to show survives: the slope is the weaker of the two forces throughout, and the shift carries the majority of the density gap in price growth under every defensible timing of supply (Section 8.4).

9 Conclusion and implications

This paper asks whether housing has become unaffordable in dense places because their supply curves are steep or because those curves move outward more slowly than in sparse locations when additional demand pressure accumulates. The distinction matters because prices do not grow when supply shifts one-for-one

with demand, however steep the curve; the shortfall is what makes prices rise, and the curve's steepness decides how much of it lands in price.

To tell the two apart we begin with the data, read through a structural model of housing supply and demand. It recovers the demand and supply shifts that observed prices and quantities imply across census tracts, and the interaction of those shifts with density gives the paper its central fact: a given percentage increase in demand buys a denser place a smaller percentage increase in supply. That fact does not rest on the model alone: a Bartik instrument applied to raw housing-unit counts finds the same density gradient with no model object on either side of the regression. Decomposing the density gap in price growth into its shift and elasticity attributions puts most of it on the shift rather than on the slope.

A theory of the buildable cap then rationalizes our findings, tying density to both objects at once: the price elasticity of supply and the rate at which the curve shifts outward. Combined with the gradient of the supply elasticity in density estimated by Baum-Snow and Han (2024), the theory lets us estimate its own primitives, and in particular the growth rate of the buildable cap as a function of density. There we find that the buildable caps of the densest neighborhoods are being raised at roughly a tenth of the pace of the least dense.

Regulation measured outside the model moved the same way. Across the Wharton survey's two waves, local land-use regulation loosened least exactly where density was high and demand was strong, which is what the recovered caps imply. And New York City's lot-level zoning files, the one place where the cap itself is on the record, show the median tract's allowed capacity unchanged across the sample window, with 25 percent of residentially zoned lots carrying buildings that the city's own zoning would not let them rebuild at their own size. What holds the cap down there is a decision rather than a geological limit: what froze was permission, not possibility.

What do our findings mean for policymakers? Doubling the densest decile's supply elasticity would lower its price growth by about 0.07 percentage points per year, whereas its annual price growth would fall by 0.30 points if supply's absorption of demand doubled. As our theory shows, the two margins are not independent, which is what makes the target easy to mistake. A location's elasticity is governed by how close it stands to its buildable cap, so relaxing the cap both moves the curve out and temporarily makes it flatter. That is why one-time reforms disappoint: a single upzoning shifts the curve once and is then spent, and as building approaches the new ceiling the elasticity falls back and price growth resumes. What dense places need in order to reduce price growth year after year is not one large relaxation but a cap that keeps moving as demand arrives, and moving faster than in sparse places, not slower.

The reforms that matter are therefore the standing ones rather than the one-off ones. By-right intensification, pre-approved building forms, standardized zoning, and time-limited permitting all change how a neighborhood responds to the *next* increment of demand, rather than granting it a fixed quantum of capacity today; redevelopment capacity in built-up areas matters more than greenfield expansion for the same reason. Such rules are likely harder to adopt exactly where they would matter most, because of the political struc-

ture of housing construction decisions: under fragmented zoning authority, the political-economy literature predicts that the places under the greatest demand pressure relax their constraints the least (Favilukis and Song, 2023; Ortalo-Magné and Prat, 2014).

One caveat follows from the same logic. Sustained permission is necessary but need not be sufficient. Where so much already stands that redeveloping it is costly, relaxing the cap frees capacity that developers may still find unprofitable to use, so the effective supply response can lag the regulatory one (Rollet, 2025). And because a given cap relaxation buys a smaller proportional shift the fuller a place already is, holding the supply response steady requires the cap to rise faster as the place fills. If there is a single lesson in these estimates, it is that affordability in dense places is not restored by one act of permission, but by rules that go on granting permission as demand continues to grow.

References

Albouy, D., Ehrlich, G., and Liu, Y. (2016). Housing demand, cost-of-living inequality, and the affordability crisis. *NBER Working Paper* 22816.

Bartik, T. J. (1991). *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research.

Baum-Snow, N., and Han, L. (2024). The microgeography of housing supply. *Journal of Political Economy*, 132(6), 1897–1946.

Brueckner, J. K., and Singh, R. (2020). Stringency of land-use regulation: Building heights in US cities. *Journal of Urban Economics*, 116, 103239.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3), 414–427.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2), 238–249.

Citizens Budget Commission (2022). *Improving New York City's Land Use Decision-Making Process*. New York, September 2022.

Damen, S., and Schildermans, S. (2022). Capital requirements, mortgage rates and house prices. *Journal of Banking and Finance*, 143, 106596.

DeFusco, A. A., and Paciorek, A. (2017). The interest rate elasticity of mortgage demand: Evidence from bunching at the conforming loan limit. *American Economic Journal: Economic Policy*, 9(1), 210–240.

Diamond, R. (2016). The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000. *American Economic Review*, 106(3), 479–524.

Einstein, K. L., Glick, D. M., and Palmer, M. (2019). *Neighborhood Defenders: Participatory Politics and America's Housing Crisis*. Cambridge University Press.

Favilukis, J., and Song, J. (2023). Why zoning is too restrictive. Working paper, SSRN 4535417.

Ganong, P., and Shoag, D. (2017). Why has regional income convergence in the U.S. declined? *Journal of Urban Economics*, 102, 76–90.

Glaeser, E. L., and Gyourko, J. (2018). The economic implications of housing supply. *Journal of Economic Perspectives*, 32(1), 3–30.

Combes, P.-P., Duranton, G., and Gobillon, L. (2021). The production function for housing: Evidence from France. *Journal of Political Economy*, 129(10), 2766–2816.

Epple, D., Gordon, B. R., and Sieg, H. (2010). A new approach to estimating the production function for housing. *American Economic Review*, 100(3), 905–924.

Glaeser, E. L., Gyourko, J., and Saks, R. (2005). Why have housing prices gone up? *American Economic Review Papers and Proceedings*, 95(2), 329–333.

Goolsbee, A., and Syverson, C. (2023). The strange and awful path of productivity in the U.S. construction sector. NBER Working Paper 30845.

Green, R. K., Malpezzi, S., and Mayo, S. K. (2005). Metropolitan-specific estimates of the price elasticity of supply of housing, and their sources. *American Economic Review Papers and Proceedings*, 95(2), 334–339.

Gyourko, J., and Molloy, R. (2015). Regulation and housing supply. In *Handbook of Regional and Urban Economics*, Vol. 5, 1289–1337. Elsevier.

Gyourko, J., Saiz, A., and Summers, A. (2008). A new measure of the local regulatory environment for housing markets. *Urban Studies*, 45(3), 693–729.

Hanushek, E. A., and Quigley, J. M. (1980). What is the price elasticity of housing demand? *Review of Economics and Statistics*, 62(3), 449–454.

Herkenhoff, K. F., Ohanian, L. E., and Prescott, E. C. (2018). Tarnishing the golden and empire states: Land-use restrictions and the U.S. economic slowdown. *Journal of Monetary Economics*, 93, 89–109.

Hilber, C. A. L., and Vermeulen, W. (2016). The impact of supply constraints on house prices in England. *Economic Journal*, 126(591), 358–405.

Hsieh, C.-T., and Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2), 1–39.

Laufer, S., and Tzur-Ilan, N. (2021). The effect of LTV-based risk weights on house prices: Evidence from an Israeli macroprudential policy. *Journal of Urban Economics*, 124, 103349.

Louie, S., Mondragon, J. A., and Wieland, J. (2026). Supply constraints do not explain house price and quantity growth across U.S. cities. Federal Reserve Bank of San Francisco Working Paper 2025-06.

Mayer, C. J., and Somerville, C. T. (2000). Residential construction: Using the urban growth model to estimate housing supply. *Journal of Urban Economics*, 48(1), 85–109.

Mayo, S. K. (1981). Theory and estimation in the economics of housing demand. *Journal of Urban Economics*, 10(1), 95–116.

Mian, A., and Sufi, A. (2011). House prices, home equity-based borrowing, and the US household leverage crisis. *American Economic Review*, 101(5), 2132–2156.

Mian, A., Rao, K., and Sufi, A. (2013). Household balance sheets, consumption, and the economic slump. *Quarterly Journal of Economics*, 128(4), 1687–1726.

Mleczko, M., and Desmond, M. (2023). Using natural language processing to construct a National Zoning and Land Use Database. *Urban Studies*, 60(13), 2564–2584.

O'Neill, M., Gualco-Nelson, G., and Biber, E. (2019). Developing policy from the ground up: Examining entitlement in the Bay Area to inform California's housing policy debates. *Hastings Environmental Law Journal*, 25(1).

Ortalo-Magné, F., and Prat, A. (2014). On the political economy of urban growth: Homeownership versus affordability. *American Economic Journal: Microeconomics*, 6(1), 154–181.

Paciorek, A. (2013). Supply constraints and housing market dynamics. *Journal of Urban Economics*, 77, 11–26.

The Pew Charitable Trusts (2024). Reforms spur faster housing approvals in California. Issue brief, August 2024.

Polinsky, A. M., and Ellwood, D. T. (1979). An empirical reconciliation of micro and grouped estimates of the demand for housing. *Review of Economics and Statistics*, 61(2), 199–205.

Rollet, V. (2025). Zoning and the dynamics of urban redevelopment. Working paper, Massachusetts Institute of Technology.

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

Song, J. (2025). The effects of residential zoning in U.S. housing markets. *Journal of Urban Economics*, 149, 103765.

U.S. Census Bureau (2025). Length of time from authorization to completion of new privately owned residential buildings, 2024. *Survey of Construction, Characteristics of New Housing*.

Tables

Table 1: Sample means and standard deviations across U.S. census tracts

Year	Real House Price	Dwellings
1999	\$189,072	1,774
	(\$127,823)	(950)
2002	\$235,870	1,846
	(\$170,432)	(1,031)
2007	\$326,355	1,980
	(\$250,016)	(1,291)
2011	\$214,912	1,999
	(\$187,966)	(1,346)
2016	\$300,288	2,016
	(\$277,321)	(1,404)
Avg. annual growth	+2.8%	+0.8%

Notes: Dwellings are total dwelling units per census tract (Section 4). Dollar values in real 2002 terms (CPI-U). Standard deviations across tracts in parentheses. Rows report selected years, and the final row the average annual growth rate of each series over the full sample. Sample: balanced annual tract panel, U.S. metropolitan areas, 1999 to 2016.

Table 2: Main regression: supply shift on demand shift and density

Variable	Coef.	Std. Err.	t	p	95% CI
(β_1) $\log \varphi$	0.4001***	0.0105	38.27	0.000	[0.3796, 0.4205]
(β_2) $\log(\text{Den}_{t-m}) \times \log \varphi$	-0.0361***	0.0017	-20.77	0.000	[−0.0395, −0.0327]
(β_3) $\log(\text{Den}_{t-m})$	0.0054***	0.0005	11.81	0.000	[0.0045, 0.0063]
Constant	-0.0505***	0.0031	-16.47	0.000	[−0.0565, −0.0445]

Notes: Dependent variable is $\log \psi$, the recovered log supply shift. Density is dwellings per square mile. HAC standard errors (panel Newey–West within tract, Bartlett kernel, two lags; see Appendix A.2 for clustered alternatives). No tract or year effects. The supply elasticity is time-varying and declines with density along the Baum-Snow and Han (2024) gradient. ***, **, and * denote significance at the 1, 5, and 10 percent levels.

Table 3: Implied supply responsiveness by density decile

Density decile	Marginal effect $\partial \log \psi / \partial \log \varphi$	Interpretation
1 (least dense)	0.261	supply closes ≈26 percent of a demand increase
10 (most dense)	0.083	supply closes ≈8 percent of a demand increase

Notes: Marginal effect equals $\beta_1 + \beta_2 \log(\text{Den}_{t-m})$ evaluated at decile-mean log density. See Figure 2 for the full profile with 95% confidence bands.

Table 4: Decomposing price growth across density deciles: shift versus slope

Density decile	Absorption ME	Supply elast. (annual)	Price growth (%/yr)	= common + shift + slope			Share of extra (%)	
				Common (pp)	Shift (pp)	Slope (pp)	Shift	Slope
1	0.262	0.077	2.51	2.51	+0.00	+0.00	–	–
2	0.207	0.058	2.77	2.51	+0.19	+0.07	72	28
3	0.179	0.048	2.91	2.51	+0.28	+0.12	71	29
4	0.161	0.042	2.99	2.51	+0.34	+0.14	70	30
5	0.148	0.037	3.06	2.51	+0.39	+0.17	70	30
6	0.138	0.034	3.11	2.51	+0.42	+0.18	70	30
7	0.129	0.031	3.16	2.51	+0.45	+0.20	69	31
8	0.121	0.028	3.20	2.51	+0.48	+0.22	69	31
9	0.109	0.024	3.26	2.51	+0.52	+0.24	69	31
10	0.082	0.014	3.41	2.51	+0.61	+0.29	68	32

Notes: Model-implied real price growth by density decile, feeding the common calibrated demand shift through each decile's supply response, $\Delta \log p = g_\varphi(1 - \text{ME})/(\varepsilon_s - \varepsilon_d)$. Absorption ME $= \beta_1 + \beta_2 \log \text{Den}$ is the share of demand growth absorbed by the supply shift (Section 5.2); the supply elasticity is the decile mean of the annual-flow elasticities. Each decile's price growth is the sum of the last three columns. The common column is the price growth of the least dense decile itself, which every decile shares: because supply absorbs only part of demand growth everywhere, prices rise even where building is easiest. The shift and slope columns then decompose the extra growth that density adds: the shift column moves absorption from the least dense decile's value to the decile's own, holding the elasticity at the least dense decile's value; the slope column then moves the elasticity to the decile's own value (the small interaction is folded into the slope column). The final two columns express the same split as shares of the decile's extra growth over the least dense decile. At every decile the shift contribution is roughly twice the slope contribution.

Table 5: The density gradient under predicted (Bartik) demand

	OLS	Reduced form	2SLS
Demand shift (recovered)	$+0.65^{***}$		$+0.78^{***}$
	(0.05)		(0.15)
Demand $\times$ log density	-0.060^{***}		-0.096^{***}
	(0.009)		(0.019)
Predicted demand B		$+1.15^{***}$	
		(0.27)	
$B\times$ log density		-0.13^{***}	
		(0.03)	
Log density	$+0.011^{***}$	-0.005	$+0.020^{**}$
	(0.004)	(0.005)	(0.008)
Response to B, least dense decile		$+0.63$	
median decile		$+0.22$	
densest decile		$+0.01$	
Tracts		19,639	
Counties		344	
Metropolitan areas		155	

Notes: Cross-section of tracts in the final sample year, in the counties (grouped into metropolitan areas) matched to the County Business Patterns. "Supply" is the tract's recovered supply shift accumulated over 1999–2016, "Demand" the recovered demand shift accumulated over the same window, both in logs as in the baseline. B is the predicted demand growth of the tract's county: its industry employment shares in the year just before the sample begins, multiplied by each industry's national log employment growth to the end of the sample, computed leaving the county's own employment out, so that no county's own boom enters its own prediction (Section 5.3).

The dependent variable is the tract's supply shift in every column. The OLS column re-estimates the baseline specification on this cross-section and sample. The reduced-form column replaces recovered demand and its density interaction with B and $B\times$ log density. The 2SLS column keeps recovered demand and its interaction as the regressors and instruments them with B and $B\times$ log density, so that only the demand variation that inheritance predicts identifies the coefficients. The response rows evaluate the reduced form, the derivative of supply with respect to B, at the mean log density of the stated decile. Because predicted employment growth is not one-for-one housing demand, the level of the response is not an absorption share; the comparison across densities is the object of interest.

Standard errors, in parentheses, are clustered by metropolitan area throughout. B now varies at the county level, so metropolitan clustering is the more conservative of the two natural choices, allowing for correlation across counties that share a labor market; clustering by county instead, the level at which B itself varies, gives tighter standard errors throughout: for example, first-stage joint F-statistics of 19.3 and 14.8 for the two endogenous regressors under county clustering, against 15.3 and 11.5 under the metropolitan clustering reported here; ***, ** and * denote significance at the 1%, 5% and 10% levels.

Table 6: Recovering the cap: implied fill fractions and cap-relaxation rates by density decile

Density decile	Recovered supply shift (%/yr)	Implied fill fraction f	Cap relaxation g (%/yr)	g under alternative fill anchors		
				0.70	0.80	0.90
1	0.25	0.45	0.112	0.09	0.13	0.18
2	0.42	0.47	0.202	0.17	0.23	0.30
3	0.33	0.52	0.175	0.15	0.19	0.25
4	0.21	0.56	0.118	0.11	0.13	0.17
5	0.14	0.59	0.086	0.07	0.09	0.11
6	0.10	0.63	0.062	0.06	0.07	0.08
7	0.06	0.65	0.039	0.04	0.04	0.05
8	0.04	0.69	0.030	0.03	0.03	0.04
9	0.02	0.71	0.011	0.01	0.01	0.01
10	0.02	0.76	0.014	0.01	0.01	0.02

Notes: Medians by decile of predetermined density. The recovered supply shift is the cumulative log ψ at the end of the sample, annualized. The fill fraction follows from the Baum-Snow and Han (2024) elasticity for each tract through the theory's supply curve, $f = 1/(1 + \alpha\varepsilon_s)$, with the cost curvature α estimated from the shape of that gradient (Section 6.5). The implied cap relaxation is $g = f \times$ shift, where the shift is the recovered supply shift in percent per year. The last three columns recompute g imposing a fill fraction of 0.70, 0.80 or 0.90 on the densest decile instead: the level of g moves with it, since a fuller location has less headroom and thus needs a larger cap relaxation to produce the same shift, while the ranking and the roughly one-tenth dense-to-sparse ratio vary very little.

Table 8: Did the cap respond? The same communities in the 2006 and 2018 Wharton surveys

Measure	2006	2018	Change, 2006–2018			Loosening, dense − sparse
			All	Sparse half	Dense half	
Approval delay index (months)	6.54	6.02	-0.52	-1.09	+0.08	-1.18
Review time, by-right single family (months)	4.00	2.43	-1.57	-1.99	-1.11	-0.88
Review time, by-right multifamily (months)	4.90	3.14	-1.76	-1.95	-1.54	-0.41
Entities that must approve a rezoning (of 5)	1.92	2.30	+0.38	+0.34	+0.42	-0.08
Entities that must approve by-right (of 6)	1.59	1.56	-0.03	-0.12	+0.07	-0.19
Share with a hard annual permit cap	0.06	0.04	-0.02	-0.05	+0.01	-0.05
Share with minimum lot size above one acre	0.27	0.34	+0.07	+0.03	+0.10	-0.07
Share with an affordable-housing requirement	0.15	0.12	-0.03	-0.04	-0.02	-0.02
Joint test: all eight loosening gaps negative; permutation $p = 0.016$ (10,000 metro-level relabelings)						

Notes: Communities that answered both waves of the Wharton land-use survey, matched by state and place code and restricted to the estimation sample's metropolitan areas (360 communities in 94 metropolitan areas). Because the aggregate index is standardized within each wave, the comparison uses the survey components measured in identical units in both questionnaires: review times in months, counts of approval entities restricted to the bodies asked about in both waves, and shares of communities with the stated rule. Sparse and dense halves split the matched communities at the median metropolitan density. The 2006 and 2018 columns report matched-sample means, and the change columns report the mean within-community change over the two waves, overall and by density half; missing items reduce the sample slightly for some rows.

The joint-test row, at the foot of the table, asks whether eight gaps sharing a sign could be chance, given that the components are correlated: metropolitan areas are reassigned to the two halves at random 10,000 times, communities inheriting their metropolitan area's label, and the reported p is the share of relabelings in which all eight loosening gaps come out negative. Because each metropolitan area moves as a whole, its communities' answers move with it, so the eight component changes remain exactly as correlated in every relabeling as they are in the data; only the split is re-randomized. Collapsing the three correlated timing measures into their average and re-running the test on the resulting six components gives $p = 0.021$.

Table 7: The cost function with a term in the standing stock

	Baseline $\theta = 0$	Estimated $\theta > 0$
Panel A. Parameters		
Curvature α	1.66	3.7
Stock exponent θ	0	0.37
Cap-density elasticity η	0.85	0.71
Fit to the shape of the elasticity gradient, R^2	0.663	0.710
Model fill, matched New York City tracts	0.81	0.65
Measured fill, matched New York City tracts	0.65	0.65
Panel B. Results of Sections 5 and 5.2		
Demand-shift coefficient, β_1	0.400	0.400
Interaction with density, β_2	-0.036	-0.036
Marginal effect, sparsest decile	0.261	0.261
Marginal effect, densest decile	0.083	0.083
Demand shift, tract mean, % a year	1.76	1.76
Supply shift, tract mean, % a year	0.32	0.32
Panel C. Cap recovery of Section 6.5		
Fill fraction, sparsest decile	0.45	0.21
Fill fraction, densest decile	0.76	0.56
Cap relaxation, sparsest decile	0.112 percent	0.072 percent
Cap relaxation, densest decile	0.014 percent	0.011 percent
Ratio, densest to sparsest	0.12	0.15

Notes: The first column is the paper's baseline, which sets $\theta = 0$ in the marginal cost (8); the second estimates α and θ jointly from the shape of the Baum-Snow and Han (2024) elasticity gradient together with the average fill fraction measured in New York City's zoning records (Section 6.6). Panel A reports the two parameters, the cap-density elasticity they imply, and how well each column meets the two targets. Panel B reports the regression and marginal effects of Sections 5 and 5.2, which are identical across the columns because the recovery of the shifts uses the Baum-Snow and Han (2024) elasticities and observed prices and quantities, never the cost function. Panel C reports the fill fractions and cap-relaxation rates each column implies, with g in percent per year.

Table 9: Did regulation loosen less in dense metropolitan areas?

	Strong-demand metros		Weak-demand metros	
Dense − sparse, loosening 2006–2018	Gap	Metros	Gap	Metros
Composite index	−0.51**	45	−0.17	45
	(0.21)		(0.20)	
Randomization p (20,000 draws)	0.026		0.390	
Continuous log density (slope)	−0.38***	45	−0.02	45
	(0.13)		(0.14)	
Halves redefined by predicted (Bartik) demand, Section 5.3				
Composite index	−0.45**	44	−0.02	44
	(0.19)		(0.24)	
Randomization p (20,000 draws)	0.020		0.948	
Continuous log density (slope)	−0.20**	44	−0.02	44
	(0.09)		(0.20)	

Notes: Each entry is the dense-minus-sparse gap in how much land-use regulation *loosened* between the 2006 and 2018 Wharton waves, estimated separately among metropolitan areas with strong and with weak demand. A negative gap means the dense group loosened less.

The composite is built the way the Wharton index itself is built (Gyourko et al., 2008). We take the seven survey components recorded in identical units in both questionnaires and observed for every matched metropolitan area: an approval-delay index and two by-right review times in months, counts of the bodies whose approval a rezoning and a by-right project need, and the shares of communities with a hard annual permit cap and with an affordable-housing requirement. The eighth comparable component, the share with a minimum lot size above one acre, is missing for some communities; including it would mean averaging seven components for some metropolitan areas and eight for others, so we leave it out, and doing the reverse changes the gap under strong demand from −0.51 to −0.48. Each component is signed so that a positive value means the rule got looser, the same direction as the recovered cap-relaxation rate g: a positive entry for a delay means months were removed, a positive entry for a share means fewer communities impose the rule. Each is then standardized across metropolitan areas and the seven are combined using the loadings of the first principal component, and the resulting index is itself standardized, so the gaps are in standard deviations. Weighting the components equally instead would give each the same say even though the three timing measures are close substitutes, correlated about 0.5 with one another, which would put three of the seven slots on one underlying thing; the factor weights settle that from the data, and they load 0.54, 0.57 and 0.60 on the timing measures against 0.16 or less in absolute value on the rest. Under strong demand the gap is negative for each of the eight comparable components taken one at a time.

Metropolitan areas are split at the median on density and, separately, at the median on the recovered demand shift; the strong half grew at about 3.4 percent per year and the weak half at about 1.0 percent, so both are growing. Density is dwellings per square mile in 2006, the first wave's year, so the split is fixed before the change being measured and is not affected by the recovery; metropolitan density is close to constant over the sample, and taking it at the end instead leaves the gaps unchanged to two decimals. The demand split uses the metropolitan area's own recovered shift over 1999–2016, the panel window; regulatory change is slow enough that the rules altered between the two waves respond to pressure that predates the first. The shift is obtained by aggregating its tracts to metropolitan totals of dwellings and land and to dwelling-weighted mean prices and elasticities and applying the recovery of Section 3 to those aggregates, which is not the same as averaging tract-level objects because the recovery is nonlinear in prices and quantities.

The randomization row reports the two-sided permutation p for the gap above it: the observed assignment of the half's metropolitan areas to the dense and sparse groups is replaced by a random assignment of the same sizes 20,000 times, and p is the share of random assignments producing a gap at least as large in absolute value, an inference that leans on no variance formula. The continuous row drops the median split and regresses the loosening composite on log density itself within each demand half; the slope is per log point of density, with heteroskedasticity-robust standard errors in parentheses.

The lower panel repeats the comparison with the strong and weak halves redefined by predicted (Bartik) demand: the demand growth implied by each metropolitan area's constituent counties' industry employment shares in the year just before the sample begins and national industry growth to the end of the sample, built at the county level and averaged up to the metropolitan area, leaving each county's own employment out of the national rates (Section 5.3). Of the metropolitan areas in the comparison, 88 are matched to the industry data.

Heteroskedasticity-robust standard errors in parentheses; ***, ** and * denote significance at the 1%, 5% and 10% levels. Stars report significance against zero, not power: at these standard errors the gap must exceed about 0.4 standard deviations to reach the 5 percent level, so an insignificant entry rules out a large gap, not a modest one.

Table 10: Robustness: year effects, elasticity specification, elasticity gradient, construction lag, and density timing

Specification	β_1 (log φ)	β_2 (interaction)	N
Baseline	0.400***	-0.0361***	299,907
With year effects	0.403***	-0.0335***	299,907
Constant ε_s per tract (2000 value)	0.304***	-0.0267***	272,720
Density-varying construction lag (m: 3–9)	0.690***	-0.0721***	276,412
Contemporaneous density (Den_t)	0.494***	-0.0494***	299,907
Population per square mile as density	0.457***	-0.0393***	299,907
Flat ε_s (gradient slope $\times$ 0)	0.333***	-0.0310***	299,892
Gradient slope $\times$ 2	0.421***	-0.0357***	300,613
Gradient slope $\times$ 3	0.383***	-0.0290***	300,995

Notes: Each row re-estimates the baseline specification under the stated change. β_1 is the demand-shift coefficient; β_2 is the coefficient on the $\log(\text{Den}_{t-m}) \times \log \varphi$ interaction. *** denotes significance at the 1 percent level.

Table 11: The estimated interaction in simulated markets with no true density gradient in the supply response

Design	True β_2	Mean $\hat{\beta}_2$	SD	95% range
A: baseline, correct specification	0	$+0.0000$	0.0002	$[-0.0005, +0.0005]$
B: flat truth, gradient assumed	0	$+0.0070$	0.0003	$[+0.0065, +0.0075]$
C: measurement error shrinking with density	0	-0.0022	0.0003	$[-0.0029, -0.0017]$

Notes: Each design simulates 300 artificial tract panels from the model of Section 3, applies the paper's recovery formulas to the simulated prices and quantities, and re-estimates the interaction regression; the table reports the distribution of $\hat{\beta}_2$ across replications. Every simulation uses the actual initial stock, initial price, density, and supply elasticity of the 24,278 sample tracts; only the shocks are drawn, with the demand-shifter drift (2 percent per year) and innovation dispersion (0.16), and the supply-shifter noise (0.10), matched to the variation in the shifts recovered from the real data. Every design sets the true supply response identically in every tract, so the true β_2 is zero. Design A uses the correct elasticities in both simulation and recovery; with the same elasticities on both sides the recovery is the exact algebraic inverse of the simulation, returning the true shifters to machine precision, so $\hat{\beta}_2 = 0$ is an identity and the row is a check on the code rather than evidence. Designs B and C are the ones the recovery gets wrong, and so the ones whose answers are not known in advance. B simulates with a flat elasticity but recovers with the estimated gradient, which leaves the recovered supply shift carrying the assumed-minus-true elasticity gap times past price growth, a term signed by density. C adds recording error to the observed housing stock, larger in sparse tracts (5 percent) than in dense ones (1 percent); because the recorded stock enters both recovered shifters, such error makes the regressor and the dependent variable share a disturbance, and grading it by density is the one error pattern that could mimic the finding. Estimates are OLS point estimates on the simulated panels (years $t \geq m$). The interaction estimated from the real data is -0.036. Significance stars are not applicable in this table: each entry summarizes the distribution of $\hat{\beta}_2$ across simulated panels rather than testing a single estimate against zero; the 95 percent range column plays the corresponding role.

Figures

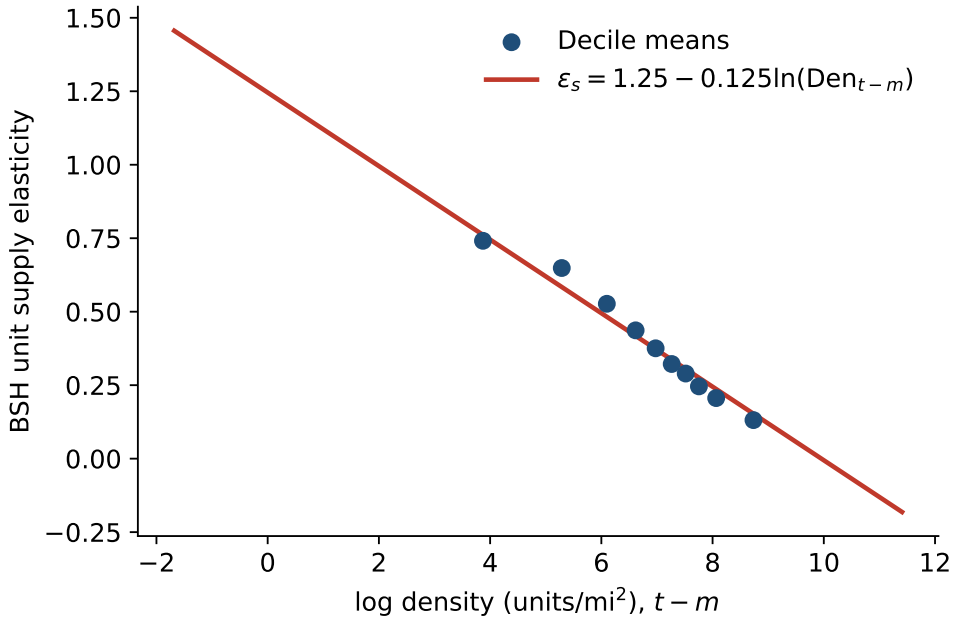


Figure 1: The density-elasticity relationship in Baum-Snow and Han (2024)

Notes: Points are mean Baum-Snow and Han (2024) unit supply elasticities by predetermined $(t-m)$ density decile; the line is the log-linear gradient fitted across tracts, with a tract-level R^2 of about 0.54. The relationship is strongly monotone.

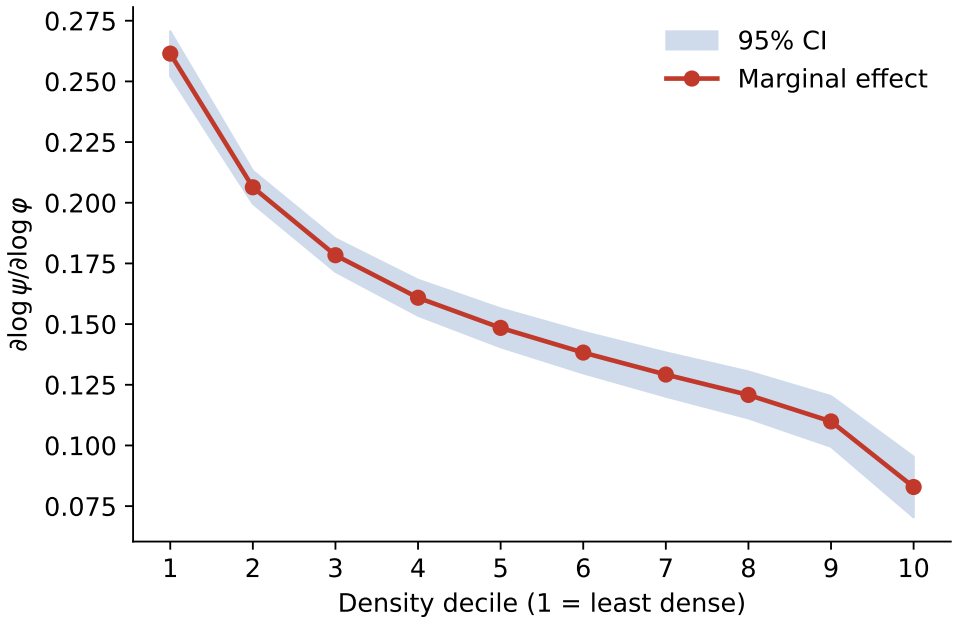


Figure 2: Supply responsiveness to demand by density decile

Notes: Marginal effect $\partial \log \psi / \partial \log \varphi = \beta_1 + \beta_2 \log(\text{Den}_{t-m})$ across density deciles, with 95% confidence bands (delta method). The response falls by about 68 percent from the least to the most dense decile.

A Additional robustness checks

Appendix A.1 discusses the implications of the cost function used in Section 6.2. The remainder consists of checks that fall into three groups. The first concerns inference: Appendix A.2 re-estimates the baseline under progressively more cautious standard errors, up to two-way clustering, and Appendix A.3 absorbs metro and metro-by-year fixed effects, so that the comparison is between neighborhoods of the same city in the same year. The second concerns the demand side: Appendix A.4 tests whether demand spilling across tract boundaries drives the interaction, Appendices A.5 and A.6 vary the demand elasticity, by assumption and with an independent measure of substitutability, and Appendix A.7 validates the recovered demand shifter against income and population growth it never saw. The third concerns the other inputs to the recovery: Appendix A.8 varies the decadal-to-annual conversion of the supply elasticity, Appendix A.9 the construction lag, Appendix A.10 splits the boom from the bust, Appendix A.11 perturbs every tract's elasticity independently within an error band, and Appendix A.12 imposes deliberately hostile measurement error on the recorded prices and quantities. The interaction of interest survives every one of these; where a check also moves the headline shift-elasticity composition, the range is reported in Section 8.4.

A.1 The cost function: why the alternatives fail

Section 6.2 builds everything on a total cost that rises as a location nears its cap and grows without bound as building approaches it. Here, we set out what the function has to deliver, whether alternatives could deliver it, what each implies relative to data we can observe, and what is left assumed once the observable implications have been used up.

What the function has to deliver. The paper's central fact is that the price elasticity of housing supply falls with density. Baum-Snow and Han (2024) estimate that gradient at the neighborhood level and we take their estimates as given. A cost function that is to be used with those estimates must be capable of producing an elasticity that varies across locations endogenously, and must tie that variation to something a location accumulates as it grows. In our function that something is fullness: $\varepsilon_s = (1/\alpha)(1-f)/f$, so a location's elasticity is a statement about how much of its cap it has used.

Constant elasticity cannot deliver it. The workhorse alternative is a cost function isoelastic in the stock, $C(H) \propto H^{1+1/\sigma}$, which is what a production function with land as a fixed factor delivers (Epple et al., 2010; Combes et al., 2021). Marginal cost is then proportional to $H^{1/\sigma}$, so it rises with building, and the supply elasticity is σ, a constant. The issue is that this elasticity does not vary with density, with fullness, or with anything else about the location.

A constant that differs across locations. One may keep constant elasticity and let the constant differ from place to place, setting σ_i exogenously. That would fit the gradient by construction. However, it would also imply that a location's elasticity is preordained and cannot change as the location grows: Los Angeles, San Francisco and New York City would have inelastic supply by virtue of where they sit, however little had been built there, and would have had it even before the first brick was ever laid. In our structure the elasticity is instead a statement about how close a place is to its cap, its fullness, which is why Section 7.3 can confront it with a measurement: New York City's measured fill is 0.65 against our inferred 0.81 for the same tracts.

A hard cap with constant elasticity predicts full locations. One may suggest the following repair: keep the constant elasticity and add a ceiling. Supply equals $\min\{(p/c)^{\sigma}, \bar{H}\}$, with ordinary marginal cost up to the limit and nothing beyond it at any price. This does generate an elasticity that falls with fullness, and if the ceiling binds lot by lot rather than tract by tract it can generate a relatively smooth gradient rather than a sharp step, if a tract's lots reach their own limits one at a time. It is therefore a natural candidate.

However, it carries an implication that can be checked against the data. Because building costs below the ceiling are bounded, sustained demand growth would push the price past the cost of the last available site, at which point the location fills its cap and stays there. High demand locations converge to $f = 1$, and aggregating over lots does not avoid this. Section 7.3 measures fullness directly for the one city where the cap is on the record. New York City's caps were essentially frozen across our window, its demand grew by 3.6 percent a year between 2011 and 2016 and 1.1 percent a year between 2004 and 2016, and yet the average tract's fill fraction was nearly constant at about 0.71 of its permitted floor area. The hard cap predicts convergence to $f = 1$; the data show a fill fraction holding steady at about 0.71.

A.2 Clustered standard errors

Our baseline specification (Table 2) reports HAC standard errors, computed as panel Newey–West within tract with a Bartlett kernel and two lags. These are robust to heteroskedasticity and to short-horizon serial correlation within a tract, but not to two other forms of dependence a tract panel can have: within-tract correlation beyond the kernel window, and, more importantly, dependence across tracts. Such dependence arises, for example, when tracts in the same metro share a common shock, such as a metro-wide price boom or bust or a regional change in labor demand, or when an aggregate housing-cycle shock hits many tracts in the same year. It makes the baseline variance assume a larger effective sample than we really have, and so understate the standard error.

We therefore recompute the baseline standard errors clustering by tract, by year, and two-way by tract and year (Cameron et al., 2011), and report all three alongside the original HAC column in Table A1. Two-way clustering is the appropriate choice given that the panel has both a cross-sectional dimension (tracts) along which shocks are likely correlated and a time dimension (years) along which aggregate housing-cycle shocks hit many tracts simultaneously. As the table shows, both coefficients survive at conventional

significance levels under every clustering scheme, so the paper's qualitative claims are unchanged. The precision the baseline column implies, however, does not survive: the clustered standard errors are several times larger, and Section 5.2 confirms that the density gradient remains distinguishable from zero under them.

A second, distinct concern is that the year dimension has only 13 clusters, too few for the asymptotic-in-G justification cluster-robust variance ordinarily relies on. We address this with a wild (Rademacher) cluster bootstrap on the year dimension, following Cameron et al. (2008), with 999 replications, imposing the null on each coefficient in turn. The bootstrap p-values, reported in the notes to Table A1, confirm that the two-way cluster result is not an artifact of the small number of year clusters.

Table A1: Standard errors under alternative clustering assumptions (baseline specification)

Standard error	HAC (baseline)	Cluster: tract	Cluster: year	Two-way (tract × year)
β_1 (log φ), coef. = 0.4001				
Standard error	0.0105	0.0153	0.0513	0.0531
t-statistic	38.27	26.18	7.80	7.53
β_2 (interaction), coef. = −0.0361				
Standard error	0.0017	0.0025	0.0048	0.0052
t-statistic	−20.77	−14.59	−7.58	−6.88
β_3 (log Den$_{t-m}$), coef. = 0.0054				
Standard error	0.0005	0.0007	0.0009	0.0011
t-statistic	11.81	7.96	5.89	4.90
N (obs. / tracts / years)	299,907 obs.; 24,278 tracts; 13 years			

Notes: Same point estimates as Table 2 (baseline specification); only the standard error changes across columns. "HAC (baseline)" is the standard error used in the main text: panel Newey–West within tract (Bartlett kernel, two lags), robust to heteroskedasticity and to serial correlation of the recovered shifters within a tract. "Cluster: tract" and "Cluster: year" are one-way cluster-robust; "Two-way" is Cameron–Gelbach–Miller multiway cluster-robust on tract and year jointly. Because 13 year-clusters is too few for the asymptotic-in-G guarantee cluster-robust variance relies on, we also report a wild (Rademacher) cluster bootstrap on the year dimension with 999 replications: the bootstrap p-value is 0.001 for both β_1 and β_2, the smallest value 999 replications can produce ($G = 13$ year clusters). A coefficient is significant at the 5 percent level when the absolute value of its t-statistic exceeds 1.96.

A.3 Metro and metro × year fixed effects

The main text omits location fixed effects because the structural recovery already nets out systematic differences across places. That argument only holds if the assigned elasticities are correctly specified for every tract; any heterogeneity they miss stays in the recovered shifters and, if it correlates with density, will load onto the interaction whether or not it reflects the mechanism we want to test. We therefore add fixed effects directly.

The choice of fixed effects is dictated by where the identifying variation lives. Only 0.4 percent of the variance of log Den$_{t-m}$ is within tract over our panel: a tract's density barely moves within the sample window, so tract fixed effects strip the density term itself of nearly all its variation. However, the interaction survives adding them too: absorbing roughly twenty-four thousand tract dummies leaves it at -0.027 ($t = -17$), and adding year effects on top changes it no further (Table A2, last row).

By contrast, 82.1 percent of the variance lies within metropolitan areas, across tracts of the same city, so metro fixed effects retain most of the identifying variation while netting out the confounds we are most worried about: regional labor demand, climate and coastal amenities, and the fact that dense tracts disproportionately sit in a particular set of high-demand metros. The panel's metro identifier assigns each tract to one of 164 metropolitan areas. Metro fixed effects compare dense and sparse tracts within the same city; metro $\times$ year fixed effects additionally net out anything specific to a city in a given year, such as a local boom.

Table A2 shows the density interaction survives both sets of fixed effects: it remains negative and significant within cities, moderately attenuated relative to the pooled baseline. We read the attenuation as informative rather than troubling: part of the pooled coefficient reflects cross-metro composition, but the larger part is a genuine within-city gradient, which is what the mechanism of Section 6 predicts, since there each location's cap responds to its own demand pressure.

Table A2: Metro and metro $\times$ year fixed effects

Specification	β_1 (log φ)	β_2 (interaction)
Baseline (pooled, two-way cluster)	0.4001^{***} (0.0531)	-0.0361^{***} (0.0052)
Metro FE (cluster: tract)	0.3912^{***} (0.0143)	-0.0295^{***} (0.0025)
Metro $\times$ Year FE (cluster: tract)	0.4036^{***} (0.0140)	-0.0234^{***} (0.0026)
Tract FE (cluster: tract)	0.2873^{***} (0.0105)	-0.0267^{***} (0.0026)

Notes: All rows use the baseline elasticities and the same estimation sample as Table 2; standard errors (in parentheses) are cluster-robust by tract in the fixed-effect rows and two-way (tract $\times$ year) in the pooled baseline row, as in the last column of Table A1. Metro fixed effects demean within each of 164 metropolitan areas; metro $\times$ year fixed effects demean within metro-year cells; tract fixed effects demean within each tract, which zeroes the intercept, so it is dropped from that row. ***, **, and * denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively.

A.4 Spatial spillovers

The estimation treats tracts as separate markets, but in reality demand can spill across their boundaries: a household priced out of a built-out tract searches next door. This raises two distinct worries, one about inference and one about bias. The inference worry is that neighboring tracts move together, so the panel carries fewer independent observations than its tract-years suggest. The first row of Table A3 concedes the point in its strongest form, clustering standard errors by metropolitan area, which allows arbitrary correlation among all of a metropolitan area's tracts across all years; the interaction remains decisive.

The bias worry is that a tract's supply shift may respond to its neighbors' demand rather than its own, and neighbors' demand may correlate with density. The remaining rows control for neighborhood pressure directly, using the final sample year's cross-section, where each tract carries its demand and supply shifts accumulated over the whole of 1999–2016. For each tract we add as a control the average of these cumulative demand shifts across the other tracts of its county, leaving the tract itself out, so the control measures the total demand pressure that built up around the tract over the sample, in the same units as the tract's own

demand regressor. The control enters negative (-0.16, $t = -7.5$), neighbors' demand pulls building away from a tract, which is the spillover operating where it should. The interaction moves from -0.062 in the same cross-section without the control to -0.052 with it, about 17 percent smaller in magnitude, and stays more negative than the panel baseline's -0.036.

The last row goes a step further and lets neighborhood pressure matter differentially by the receiving tract's density, adding the control's interaction with log density; the tract's own interaction moves only to -0.048, still more negative than the panel baseline's, so cross-tract displacement, even density-dependent displacement, is not what it is picking up.

Table A3: Spatial-spillover checks

	β_1 (log φ)	β_2 (interaction)	Obs.
Panel baseline, clustered by metropolitan area	0.4001^{***}	-0.0361^{***}	299,907
	(0.0270)	(0.0046)	
Final-year cross-section, same clustering	0.6524^{***}	-0.0622^{***}	20,020
	(0.0440)	(0.0073)	
Adding the county leave-one-out demand control	0.6514^{***}	-0.0518^{***}	20,020
	(0.0390)	(0.0061)	
Adding the control's density interaction	0.6281^{***}	-0.0481^{***}	20,020
	(0.0460)	(0.0075)	

Notes: The first row is the baseline specification and sample of Table 2 with standard errors clustered by metropolitan area. The second row estimates the same specification using the final sample year's cross-section of cumulative shifts, clustered by metropolitan area; the third adds the county leave-one-out demand control, the mean cumulative recovered demand shift of the other tracts in the tract's county (single-tract counties drop). Standard errors in parentheses; ***, ** and * denote significance at the 1 percent, 5 percent and 10 percent levels.

A.5 Varying the demand elasticity

Section 3.5 shows that the demand elasticity cancels out of the recovered supply shift and enters the regression only through the demand-shift regressor. A different value does not simply rescale that regressor: it adds a multiple of cumulative price growth, which differs across tracts and systematically with density. A pure rescaling could move only the magnitude of the coefficients; this change carries no such algebraic guarantee for the sign, so the sign has to be checked rather than assumed. In the data, only the magnitude moves. Section 5.3 checks the sign without any demand elasticity: there the demand-shift regressor is replaced outright by demand predicted from pre-sample industry mix, so no demand elasticity enters the specification at all, and the density gradient comes out stronger than in the baseline. This appendix reports the variants for completeness.

Table A4 sets the elasticity to $\varepsilon_d = -0.3$ and $\varepsilon_d = -1.0$, a range that brackets the estimates of the demand literature, and then lets it vary with income, declining in absolute value from -1.0 in the lowest-income tracts to -0.3 in the highest, on the view that housing is a smaller share of richer households' budgets. We assign the elasticity by a tract's position in the income distribution, linearly between those two endpoints, so the schedule covers that whole range and its sample average is exactly the baseline -0.65. The interaction

retains its sign and significance in all three.

Table A4: Robustness: the demand elasticity

Specification	β_1 (log φ)	β_2 (interaction)	N
Baseline	0.400***	-0.0361***	299,907
$\varepsilon_d = -0.3$	0.499***	-0.0284***	299,907
$\varepsilon_d = -1.0$	0.260***	-0.0252***	299,907
Income-dependent ε_d (-1.0 to -0.3)	0.449***	-0.0444***	299,907

Notes: Each row re-estimates the baseline specification under the stated demand elasticity. β_1 is the demand-shift coefficient; β_2 is the coefficient on the $\log(\text{Den}_{t-m}) \times \log \varphi$ interaction. The income-dependent row assigns each tract an elasticity by its position in the income distribution, linearly between -1.0 and -0.3, so that the sample average is the baseline -0.65. *** denotes significance at the 1 percent level.

Income, rather than density, is the right variable to drive the heterogeneity in the demand elasticity. The demand elasticity describes household behavior, which is governed by preferences and budgets, and in household data the budget share of housing falls with income. Income also captures what density cannot, namely the amenities that make a location worth paying for: a public-housing tower block is among the densest environments in any city, yet demand to live there is elastic rather than inelastic, precisely because the surrounding amenities are poor. An elasticity declining mechanically with density would get such places exactly wrong.

A.6 Demand elasticity varying with tract substitutability

The price elasticity of demand for any good rises with the number and closeness of its substitutes: a good with many near-identical alternatives faces a nearly horizontal demand curve, because a small price increase is enough to send buyers elsewhere, while a good with no substitutes can absorb a much larger price change before quantity demanded responds. Housing is no exception. A household facing a price increase in a tract that closely resembles several other tracts nearby, in what it costs, who lives there, and what it offers, can relocate to one of them at little cost; a household in an idiosyncratic, one-of-a-kind neighborhood has nowhere equivalent to go, and so continues to demand housing there even as its price rises.

The baseline specification treats the demand elasticity as a single national constant. This appendix asks what happens to our headline interaction if that constant is instead allowed to vary tract by tract with an independent measure of how substitutable each tract's housing actually is, in a similar spirit to the income-dependent schedule of Appendix A.5.

We build the substitutability score from a cross-section, one observation per tract. For each tract we standardize seven characteristics: log density, log house price, log household income, distance to the central business district, the tract's own supply elasticity from Baum-Snow and Han (2024), the tract's share of non-Hispanic White residents, and its share of adults aged 25 and over holding a bachelor's degree or higher. Let $z_i \in \mathbb{R}^7$ denote tract i's vector of standardized characteristics and $\hat{\Sigma}$ their covariance matrix. For any two

tracts i, j in the same metropolitan area, define the Mahalanobis distance

$$d_{ij} = \sqrt{(z_i - z_j)' \, \hat{\Sigma}^{-1} \, (z_i - z_j)}, \tag{25}$$

which reduces to ordinary Euclidean distance if the characteristics are uncorrelated, and otherwise down-weights differences along directions the characteristics already move together on, so that two highly correlated dimensions (density and price, say) are not counted twice as independent sources of dissimilarity. For each tract i let $N_K(i)$ be the set of its $K = 5$ nearest neighbors within its own metropolitan area under (25), and define the substitutability score as

$$\text{score}_i = \frac{1}{1 + \frac{1}{K} \sum_{j \in N_K(i)} d_{ij}}, \tag{26}$$

a number between zero and one that rises toward one as a tract's nearest comparators become characteristically identical to it ($d_{ij} \to 0$) and falls toward zero as even its closest comparators sit far away in characteristic space. What makes the score a measure of substitutability rather than mere similarity is the choice of characteristics. Two tracts close in (25) look alike on exactly what a household weighs in choosing where to live, namely cost, how built up the tract is, the income, racial and educational composition of its residents, distance to jobs, and how tightly its supply is constrained, so a household indifferent on those margins should regard them as close substitutes.

We restrict the comparison set in (25)–(26) to tracts within the same metropolitan area: a household choosing a neighborhood is choosing among tracts it could plausibly live in and commute from, not among tracts anywhere in the country, so only a within-metro comparator is a genuine substitute. This restriction is also what gives the score any traction. Had we instead computed it across metropolitan areas, the comparison set for every tract would jump from a few dozen or a few hundred local alternatives to the full national sample of tens of thousands of tracts, and at that scale nearly every tract would find at least one near-identical match somewhere in the country by sheer force of numbers, whatever its true local substitutability. The score itself is not monotone in density: it rises through roughly the eighth density decile before falling back in the two densest deciles, consistent with the idea that the densest urban cores are characteristically distinctive even relative to other dense tracts in the same city.

We assign each scored tract's demand elasticity by its percentile rank in the national distribution of the score, running linearly from $\varepsilon_d = -0.3$ at the least substitutable tract to $\varepsilon_d = -1.0$ at the most substitutable, exactly the rank-linear construction used for the income-dependent schedule of Appendix A.5: because ranks are uniform, this reallocates elasticity across tracts without moving its average, which therefore matches the constant baseline of -0.65. Tracts without a score, those in single-tract metropolitan areas or lacking the demographic data, keep the baseline -0.65. The bounds -0.3 and -1.0 are not fitted; they are chosen to bracket the estimates of the household demand literature surveyed in Section 4.1, so the exercise asks how

the result moves under a defensible allocation of that range rather than under an arbitrary one.

Table A5 reports the re-estimated specification: the interaction rises in magnitude from -0.036 at the constant baseline to -0.0532, an increase of about 47 percent, and remains significant under two-way clustering ($t = -8.23$, against $t = -6.88$ at baseline). The level term β_1 and the standalone density term β_3 move in the same direction, from 0.40 to 0.5152 and from 0.005 to 0.0073 respectively, so nothing in the sign or the statistical significance of the paper's central estimates depends on holding the demand elasticity fixed across tracts.

Widening the bounds well beyond the literature's supported range, to $\varepsilon_d \in [-0.1, -2.0]$ in place of $[-0.3, -1.0]$, leaves the sign and significance intact but moderates the magnitude: the interaction comes to -0.0392 ($t = -7.12$ under two-way clustering), between the constant-elasticity baseline of -0.036 and the $[-0.3, -1.0]$ result of -0.0532. In this wider variant β_1 and β_3 fall below their baseline values, to 0.3455 and 0.0020 respectively, and both remain significant.

Table A5: Demand elasticity varying with tract substitutability

Standard error	HC1	Two-way (tract × year)
β_1 (log φ), coef. = 0.5152		
Standard error	0.0068	0.0632
t-statistic	75.25	8.16
β_2 (interaction), coef. = -0.0532		
Standard error	0.0011	0.0065
t-statistic	-48.98	-8.23
β_3 (log Den$_{t-m}$), coef. = 0.0073		
Standard error	0.0003	0.0011
t-statistic	26.65	6.93
N (obs. / tracts / years)	299,907 obs.; 24,278 tracts; 13 years	

Notes: Same specification and estimation sample as Table 2, except the constant demand elasticity ($\varepsilon_d = -0.65$) is replaced by a tract-specific ε_d running linearly in the tract's national rank in the substitutability score, which is computed against comparators in the same metropolitan area, from -0.3 (least substitutable) to -1.0 (most substitutable); tracts without a score keep -0.65. See Appendix A.6 for the construction. Standard errors: HC1 (heteroskedasticity-robust) and Cameron–Gelbach–Miller two-way clustered by tract and year, following Appendix A.2. N matches the baseline exactly.

A.7 Validating the recovered local demand shifter

The demand elasticity is taken from the household demand literature (Section 4.1), and applied uniformly to every tract. That leaves open whether the resulting tract-level recovered demand shifter is a real measure of local demand pressure or just an algebraic residual. We check this by asking whether the shifter co-moves with observable demand fundamentals that played no part in the recovery. For each tract we compute the ten-year growth in the recovered demand shifter, $\Delta \log(1 + \varphi)$, over 2000 to 2010, and regress it on fundamentals measured over the same window: the growth in real household income, taken directly from the panel, and the growth in census tract population. This window is exactly the span between the two

Censuses in which tract income is genuinely observed (Section 4.1), so the fundamentals are measured, not interpolated, and the elasticity was taken from outside our data altogether, so no moment of ours, aggregate or cross-tract, entered it.

Table A6 reports the results; tract population comes from the Baum-Snow and Han (2024) construction file. Income growth and population growth each predict demand-shifter growth strongly, positively, and significantly on their own, with coefficients of roughly 0.8; entered jointly, both remain large and significant, indicating that they carry complementary information about where demand pressure arrived. Population growth alone accounts for about thirty-five percent of the cross-tract variation in demand-shifter growth. The recovered shifter, in short, rises where people and purchasing power actually went. Since neither fundamental was used anywhere in the recovery, this is evidence that the structural recovery of Section 3 separates demand from supply in a way consistent with independent data, rather than mechanically fitting prices and quantities.

Table A6: Recovered demand shifter growth against income and population growth, 2000–2010

Demand shifter growth, $\Delta \log(1+\varphi)$, 2000–2010	(1)	(2)	(3)
Real income growth	0.815***		0.675***
	(0.014)		(0.012)
Population growth		0.847***	0.785***
		(0.012)	(0.011)
Constant	0.098	0.043	0.061
R^2	0.170	0.354	0.468
N (tracts)	24,112	24,112	24,112

Notes: Tract-level cross-section. All variables are ten-year log changes, 2000 to 2010, the window between the two Censuses in which tract income is observed. Demand shifter growth is $\Delta \log(1+\varphi)$ from the model recovery (Section 3); income growth is real (CPI-U deflated) average household income from the panel; population growth is the change in log census tract population, 2000 to 2010, from Baum-Snow and Han (2024). Heteroskedasticity-robust standard errors in parentheses (the regression is a single cross-section, so there is no serial dimension for a HAC correction). ***, **, and * denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively.

A.8 The decadal-to-annual elasticity conversion

Section 3.3 converts the Baum-Snow and Han (2024) decadal elasticity into an annual-flow elasticity by dividing it by ten, since ten years of cumulative construction is roughly the sum of ten annual flows. This is a first-order approximation, so we check how much it matters by re-running the baseline with divisors of 1 (no conversion at all), 5, 8, 12, and 15 in place of 10. Table A7 shows that both coefficients move smoothly and monotonically with the divisor: a smaller divisor implies a larger annual elasticity and scales β_1 and β_2 up in magnitude. The sign and significance of both are unaffected across the entire range. The conversion factor is, as claimed in Section 3.5, a magnitude assumption, not one that can overturn the qualitative result.

Table A7: Sensitivity to the decadal-to-annual elasticity rescaling factor

Rescaling factor	Median annual ε_s	β_1 (log φ)	β_2 (interaction)
1 (no conversion)	0.356	0.953^{***}	-0.0884^{***}
1/5	0.071	0.462^{***}	-0.0419^{***}
1/8	0.044	0.415^{***}	-0.0376^{***}
1/10 (baseline)	0.036	0.400^{***}	-0.0361^{***}
1/12	0.030	0.390^{***}	-0.0351^{***}
1/15	0.024	0.380^{***}	-0.0342^{***}

Notes: The baseline divides the Baum-Snow and Han (2024) decadal elasticity by 10 (Section 3.3); this table instead divides by 1 (no conversion), 5, 8, 12, and 15, holding everything else fixed. HAC standard errors. ***, **, and * denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. The sign and significance of both coefficients are unaffected; only their magnitude moves with the assumed conversion factor.

A.9 Construction-lag mechanics

Section 8.1 shows that letting the construction lag rise with density, from three years in the least dense decile to nine in the densest, leaves the density interaction negative and significant. But the estimates also move with the lag schedule in a way worth a closer look, since the length of the lag matters on its own, not just its link to density.

Table A8 isolates this by re-running the recovery with a common (flat, non-density-varying) lag of three, four, and five years, alongside the density-varying schedule. Both coefficients scale smoothly with the assumed common lag, larger at shorter lags, because a shorter lag ties the recovered supply shift to more recent prices and attributes more of the observed quantity response to the current demand episode. Across the entire range, however, the interaction retains its negative sign and statistical significance. Two conclusions follow. First, the direction of the density dampening does not depend on the assumed construction delay within the empirically plausible range. Second, the magnitudes in the density-varying row of Table 10 partly reflect the shorter lags that the schedule assigns to less dense tracts, not only the density mechanism itself, so we treat the baseline common five-year lag as the cleaner specification for magnitudes and the density-varying row as a directional check.

Where the three-to-nine-year schedule comes from. The lag in the model is the time between the demand a builder acts on and the completed homes, so the right external benchmark is not construction time alone but the whole pipeline of approval, permitting and construction, each stage of which is documented separately.

In the Census Bureau's Survey of Construction, a new single-family house takes about nine months from permit to completion (U.S. Census Bureau, 2025); in the sparsest deciles, where approval is close to automatic, the pipeline is essentially the build itself, and a three-year lag is generous. Multifamily construction is slower, about twenty months on average and about two years for buildings of twenty or more units (U.S. Census Bureau, 2025).

In the densest, most heavily regulated markets, the binding stage sits before the permit: the median

multifamily project in San Francisco spends about two and a quarter years in planning entitlement alone, before any building permit is issued (O'Neill et al., 2019; Pew Charitable Trusts, 2024), and the median New York City project that needs a rezoning spends about two and a half years from application to the end of public review (Citizens Budget Commission, 2022), with large and contested projects in the tails of both distributions running well past either median. For a large project in a dense market, two or more years of approval, the permitting that follows, and about two years of construction together put the complete pipeline at five years and beyond, with mega-projects longer still. A lag that rises from three years in the least dense decile to nine in the densest spans these documented timelines rather than stretching them.

Table A8: Isolating the mechanical effect of construction-lag length from its density-dependence

Lag schedule	β_1 (log φ)	β_2 (interaction)
Flat lag $m = 3$ yrs	0.697^{***}	-0.0683^{***}
Flat lag $m = 4$ yrs	0.538^{***}	-0.0508^{***}
Flat lag $m = 5$ yrs (baseline)	0.400^{***}	-0.0361^{***}
Density-varying lag (3–9 yrs)	0.690^{***}	-0.0721^{***}

Notes: The first three rows apply a common (non-density-varying) construction lag of the stated length to every tract; the last row is the density-varying schedule of Section 8.1 (3 years in the least dense decile to 9 in the densest). HAC standard errors. ***, **, and * denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. A shorter common lag ties the recovered supply shift to more recent prices, scaling both coefficients up, and also retains more early-year observations; the interaction is negative and significant at every lag length shown. See Section 8.1 for discussion.

A.10 Boom versus bust

A separate question is whether the pooled result is an artifact of pooling the 2000s boom with the post-2008 bust, given how sharply real prices in Table 1 rise and then fall over the sample. Table A9 splits the sample into the boom (2004–2007) and the bust and recovery (2008–2016) and re-estimates the baseline separately in each. The interaction is negative and significant in both sub-periods.

Table A9: Boom versus bust: is the density interaction a cyclical artifact?

Subsample	β_1 (log φ)	β_2 (interaction)
Full sample (2004–2016)	0.400^{***} (0.0153)	-0.036^{***} (0.0025)
Boom (2004–2007)	0.235^{***} (0.0092)	-0.023^{***} (0.0013)
Bust & recovery (2008–2016)	0.530^{***} (0.0197)	-0.048^{***} (0.0032)

Notes: "Boom" restricts the sample to 2004–2007 (the real-price run-up preceding the 2007 peak in Table 1); "Bust & recovery" restricts to 2008–2016. Standard errors (in parentheses) are cluster-robust by tract. ***, **, and * denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. The interaction β_2 is negative and significant in both sub-periods, so the density-dampening result does not appear to be an artifact of pooling the 2000s boom with the post-2008 bust and recovery.

A.11 Imprecision in individual tract elasticities

The stress test of Section 8.2 tilts the whole elasticity schedule at once. The complementary worry is idiosyncratic: Baum-Snow and Han (2024) publish point estimates but no tract-level confidence intervals, so each individual tract's elasticity is measured with unknown precision. We therefore treat the estimates as noisy and ask how far our coefficients can move. In each replication of a simulation exercise, every tract's elasticity is perturbed by an independent draw from a ± 0.5 band, in the decadal units of the original estimates, and the perturbed values are clipped to the support of those estimates. A band of ± 0.5 is deliberately extreme: the estimates themselves mostly lie between 0.2 and 0.7, so each draw allows a tract's true elasticity to sit anywhere from zero to well above the sample maximum. We then redo the recovery of the supply shifter and the estimation from scratch with the perturbed elasticities, and repeat this one hundred thousand times.

The result is unambiguous: the density interaction is negative in every one of the hundred thousand replications, with a 95 percent range of $[-0.037, -0.035]$ around the baseline -0.036; β_1 and β_3 are similarly stable. The reason the estimates barely move is instructive. Independent errors across roughly twenty-four thousand tracts largely average out in a regression; what they add is noise in the recovered supply shifter, which costs precision but cannot manufacture or destroy a systematic pattern across density. The dangerous form of elasticity error is instead the systematic one, an entire gradient tilted up or down, and that is precisely what the stress test in the main text already ruled out. Together the two checks cover both ends: the gradient can be wrong wholesale, or every tract's elasticity can be individually imprecise, and the finding survives either way.

A.12 Mismeasured prices and quantities

The checks so far ask whether the elasticities could be wrong. The remaining inputs are the observed prices and quantities themselves, and they deserve the same scrutiny, for a structural reason: the recovery builds both shifters from the same two observed series, so an error in a tract's recorded price or stock enters the demand measure and the supply measure at the same time. Tract price indices are estimated from finite transactions and are noisiest where markets are thin, and recorded stocks come from assessor files with documented breaks (Section 4). We therefore ask two questions: could input error create our finding where none exists, and does it threaten the estimate we have?

Table A10, Panel A, answers the first question on simulated data: we simulate housing markets in which supply responds to demand identically at every density, so the true interaction is zero by construction, and we let their prices and quantities be observed with error before applying the recovery. Classical price error, even at five percent everywhere, produces an interaction indistinguishable from zero. Price error that shrinks with density, the empirically realistic pattern since sparse, thin markets have the noisiest indices, produces a small *positive* interaction: realistic price error biases against our finding, not toward it. Even the deliberately hostile design, price and quantity error together, correlated with each other and concentrated

in sparse tracts, produces only about -0.002, less than a tenth of our estimate of β_2 in Section 5.

Panel B answers the second question by adding error to the real data itself. We inject extra classical noise, independent across tract-years and uncorrelated with density, into the actual panel's log prices and log quantities, and re-run the recovery and the regression at each dose, using a replication of the pipeline that reproduces the baseline interaction on the clean data. Density-graded error patterns are covered by Panel A. At a noise standard deviation of one or two percent, already larger than plausible index error, the interaction moves by about one percent of its value or less; at an implausibly large ten percent it attenuates by about 19 percent and remains negative, with the entire range of replications far from zero, so it could bias our results in Section 5 only negligibly. If mismeasurement were generating the result, making the data worse on purpose would move it substantially; even under extreme assumptions, it does not.

Table A10: Can mismeasured prices and quantities explain the interaction?

Panel A: simulated null worlds (true $\beta_2 = 0$), corrupted observation				
	True β_2	Mean $\hat{\beta}_2$	SD	95% range
E: classical price error (5% everywhere)	0	$+0.0001$	0.0003	$[-0.0005, +0.0006]$
F1: price error falling with density (5% $\rightarrow$ 1%)	0	$+0.0005$	0.0002	$[-0.0000, +0.0009]$
F2: price error rising with density (1% $\rightarrow$ 5%)	0	-0.0004	0.0002	$[-0.0008, +0.0000]$
G: joint price and quantity error ($\rho = 0.7$)	0	-0.0022	0.0003	$[-0.0027, -0.0017]$
Panel B: extra noise injected into the real panel (both series)				
	Noise s.d.	Mean $\hat{\beta}_2$	SD	95% range
No added noise (baseline replication)	–	-0.0360	–	–
Injected error	0.01	-0.0359	0.0001	$[-0.0361, -0.0357]$
Injected error	0.02	-0.0356	0.0002	$[-0.0361, -0.0353]$
Injected error	0.05	-0.0341	0.0006	$[-0.0352, -0.0331]$
Injected error	0.10	-0.0293	0.0010	$[-0.0308, -0.0272]$

Notes: Panel A simulates artificial tract panels with corrupted observation: each of 300 replications simulates a tract panel with no true density gradient in supply behavior (true $\beta_2 = 0$), corrupts the observed log prices and/or log quantities as stated, and applies the paper's recovery and regression. Design E adds classical price noise (s.d. 5 percent) to every tract-year; F1 grades the price noise from 5 percent in the sparsest tracts to 1 percent in the densest (the realistic direction, since thin markets have the noisiest indices); F2 reverses the profile; G corrupts prices and quantities together on the F1 profile with correlation 0.7 within tract-year. Panel B injects extra classical noise of the stated standard deviation into the actual panel's log prices and log quantities (100 replications per dose) and re-runs the recovery and regression, using a replication of the pipeline whose clean-data run reproduces the baseline interaction. The observed baseline estimate is -0.036.

A.13 Reconciling the debate on supply constraints

Our work is also related to Louie et al. (2026). Comparing metropolitan areas across four leading measures of supply constraint, they find that income growth predicts essentially the same house-price growth and the same housing-quantity growth whether a city is measured as constrained or not, and conclude that measured supply constraints do not explain the growth of housing prices and quantities. In this section we argue that their findings are not evidence against supply constraints.

Their first result is about price growth: across metropolitan areas, measured elasticity does not predict how strongly income growth passes into house prices. Two distinct things produce that null, and separating

them is what reconciles their evidence with ours. The first is scale. The slope does govern capitalization: in equilibrium, price growth is unabsorbed demand growth divided by the sum of the absolute price elasticities of supply and demand, so the elasticity sits in the denominator and must show up somewhere. It shows up across census tracts, and it vanishes across metropolitan areas, where an elasticity averaged over neighborhoods standing at their caps and neighborhoods with ample room is a good proxy for neither. The second is magnitude. Even measured where it binds, the slope is the smaller channel: our counterfactuals (Section 5.2) attribute about 29 percent of the density gap in price growth to it, against 71 percent to the shift. Sorting cities on measured elasticity is therefore both a noisy way to capture the constraint and a search on the margin that does less of the work, and their null follows on both counts.

Their second result is about quantities: measured elasticity does not predict how much extra housing income growth calls forth. This follows for the same reason we have argued throughout the paper: if building happens when caps are raised rather than when demand pushes developers up a supply curve, then demand pressure and measured slope have no reason to interact in the quantity data either. Section 5.2 gave the magnitudes behind this claim. Had supply curves never shifted out, the least dense decile would have added about 0.26 percent of housing per year instead of the 1.5 percent it did add, while the densest would have added about 0.07 percent against the 0.0 percent it did add: where the cap does not move, there is little building for either account to explain.

Their own analysis runs into the same requirement ours does: producing the observed quantities takes supply shifts. It surfaces in two places. First, in an appendix titled "Some Unpleasant Housing Supply Arithmetic," they show that their estimates require supply shifts several times more important than income growth. They call the result unpleasant because they are "not aware of work that argues for a prominent role of housing supply shocks (as opposed to different elasticities)." Second, explaining why less-constrained cities saw lower price growth, they conclude that "without these shifts in housing supply, overall growth in housing quantities would have been essentially zero in less-constrained cities." Their own arithmetic already calls for a shift; what they lacked was a credible mechanism.

Despite needing a shift, they reject the supply-shift explanation, giving two main reasons. The first is that the implied shifts "are far too large to be plausible, or at least they would imply enormous improvements in construction productivity that have not been documented." This is decisive against their supply shift but not against ours, because in their model a supply shift is a gain in construction productivity, while the shift we recover is the buildable cap moving: rezonings, approvals, height and density limits.

What separates the two readings is not the size of the shift but what it would have to represent. Read as productivity, the shift explanation runs against a measured record: Goolsbee and Syverson (2023) document that labor productivity and total factor productivity in U.S. construction have fallen more or less continuously since 1970, so a construction worker in 2020 produced less than one in 1970. Any positive figure contradicts that record, and the objection stands however small it is. Read as zoning, it contradicts nothing, because tract-level buildable capacity is not measured at all, and a drift of the size we recover is in fact quite

feasible.

The size of the cap relaxation is small in cumulative terms. The cap expands by about 0.1 percent to 0.2 percent a year in the least dense deciles and about 0.01 percent in the densest (Section 6.5), which over seventeen years is two to three and a half percent and a quarter of one percent respectively. For the densest decile, whose 3.4 million dwellings sit under caps allowing about 4.7 million, that is room for some eleven thousand additional homes. Undocumented technical progress is implausible; regulatory capacity being released at that pace is not. Section 7 showed that independently measured regulation moved in the same direction as these recovered caps imply.

The second reason they give for rejecting the shift is more subtle, and worth stating carefully. If supply shifts are to explain both of their findings at once, the shifts must behave in two opposite ways. To produce the first pattern (the tendency of less-constrained cities to show lower price growth), shifts must be larger where measured elasticity is higher. But to produce the absence of any differential response to income growth that they find, shifts that move with income must be smaller where elasticity is higher, since otherwise elastic cities would visibly convert more of their income growth into quantities. They call this combination epicyclic and observe that no theory of construction productivity would deliver opposite correlations depending on whether income happens to be growing.

Our account needs no such combination, because the second of these patterns, the absence of a differential price-growth response to income growth, disappears once we go down to the census tract. That leaves only the first pattern (the tendency of less-constrained cities to show lower price growth) for the shift to account for, and shifts that are larger where there is room to build (the elastic geographies) explain it, exactly as their own derivation requires.

The disappearance of the second pattern can be shown directly. Table A11 runs their regression in our data at both scales, using only raw objects: tract price and housing-stock growth over 1999–2016, growth in average household income between the 2000 and 2010 Censuses, and the Baum-Snow and Han (2024) elasticity as shipped. Across metropolitan areas we reproduce their price null result ($t = -0.8$). Across census tracts, income growth passes into price growth at a rate of about 0.86 where measured elasticity is low but only about 0.43 where it is high ($t = -3.6$), and comparing tracts within the same metropolitan area barely changes this. The two facts fit together. Pass-through measures the share of a given demand increment that becomes price; the density gap depends also on how large that increment is, and it is there that dense and sparse neighborhoods differ most. The slope is real and does what the theory says. It is simply not the main driver.

Their quantity null result, by contrast, holds at every scale including ours, exactly as the counterfactual above implies.[41]

Their evidence and ours describe the same world. Across metropolitan areas, measured elasticity predicts neither price growth nor quantity growth, and we find that too. What we add is why. The elasticity is

[41] This exercise follows Louie et al. (2026) in style rather than replicating them: they study different horizons, with additional controls and other elasticity measures.

not missing from the data but averaged away, because the constraint binds neighborhood by neighborhood and a city-wide elasticity blends neighborhoods standing closer to their caps with neighborhoods that still have room. The reallocation this induces, demand deflected from constrained neighborhoods into their less constrained neighbors, is netted out by the same aggregation. If demand crosses tract lines this way, our own tract-level estimates could inherit the problem, with a tract's supply answering its neighbors' demand rather than its own. Appendix A.4 showed that controlling for neighbors' demand leaves the interaction intact. And even where it does bind, the slope is the smaller channel, because what governs how much demand goes unabsorbed is mostly the pace at which buildable caps are expanded. This suggests not that supply elasticities fail to matter, but that the constraint lives mostly in the shift.

Table A11: The elasticity regression at two scales, in prices and in housing stock

	Census tracts	Census tracts	Metropolitan areas
Panel A: price growth			
Income growth	$+0.959^{***}$	$+0.625^{***}$	$+2.337^{**}$
	(0.149)	(0.076)	(1.039)
Income growth × elasticity	-0.727^{***}	-0.573^{***}	-1.677
	(0.200)	(0.125)	(2.070)
Elasticity	$+0.013$	$+0.018^{***}$	$+0.041$
	(0.010)	(0.005)	(0.051)
R^2	0.18	0.12 (within)	0.33
Panel B: housing-stock growth			
Income growth	$+0.064^{***}$	$+0.055^{**}$	-0.617
	(0.019)	(0.024)	(0.414)
Income growth × elasticity	$+0.055$	$+0.114$	$+1.139$
	(0.068)	(0.071)	(0.786)
Elasticity	$+0.023^{***}$	$+0.022^{***}$	-0.016
	(0.004)	(0.003)	(0.020)
R^2	0.13	0.13 (within)	0.05
Metro fixed effects	No	Yes	—
Observations	18,805	18,805	122

Notes: The dependent variable is annualized log house-price growth over 1999–2016 in Panel A and annualized log housing-stock growth over the same window in Panel B. Income growth is the annualized log growth of average household income between the 2000 and 2010 Censuses; the elasticity is the tract-level measured supply elasticity of Baum-Snow and Han (2024), exactly as shipped with the data. No object recovered by our model enters the table. Columns 1 and 2 are estimated on census tracts, column 2 within metropolitan areas (its R^2 is the within fit); column 3 is estimated on metropolitan-area means for the 122 metropolitan areas with at least 20 sample tracts, the unit of observation in Louie et al. (2026). Standard errors (in parentheses) are clustered by metropolitan area in columns 1–2 and heteroskedasticity-robust in column 3. ***, **, and * denote significance at the 1, 5, and 10 percent levels.