Highly Disaggregated Land Unavailability

Chandler Lutz
University of North Carolina, Charlotte

Ben Sand
York University

Land Unavailability Data:
https://github.com/ChandlerLutz/LandUnavailabilityData

January 11, 2026

Abstract

Standard empirical proxies contradict the canonical prediction that supply constraints amplify house price cycles. We resolve this puzzle by constructing the Land Unavailability–Machine Learning (LU-ML) Indices using high-resolution satellite imagery. Unlike existing measures, our indices capture the nonlinear and heterogeneous impacts of physical geography. We document two main facts: (1) physical constraints are a key determinant of cross-sectional price dynamics, driven largely by the intensive margin (steep slopes); and (2) looser supply constraints significantly mitigate the price effects of demand growth, overturning the "null result" in recent literature.

JEL Classification: R30, R31, R20;

Keywords: Land Unavailability, Supply Constraints, House Prices, Entrepreneurship

Amid a crisis of U.S. home affordability, households and legislators have targeted supply constraints as both a culprit for high prices and a lever for policy response (The White House, 2025).[1] This focus matches economic theory: building constraints steepen the housing supply curve, amplifying the price impact of demand shocks. Yet standard empirical proxies for these constraints have limited predictive power for house prices (Davidoff, 2013, Guren et al., 2021, Figures 1 and 4), while lower constraints, according to these measures, *do not* mitigate the house price impacts of demand shocks (Louie et al., 2025, Figures 2 and 8). These results, taken at face value, have wide-reaching ramifications for housing markets and upend urban theory, implying that (1) supply constraints explain little of the heterogeneous price growth across cities; (2) demand shocks have similar price effects regardless of supply elasticities; and (3) the extensive literature that exploits supply constraints to sort housing markets or for exogenous variation is unfounded.[2]

This paper resolves this disconnect between theory and evidence by developing novel supply constraint measures that combine new Land Unavailability (LU) estimates from highly disaggregated satellite imagery with machine learning (ML) techniques. The resulting LU-ML indices capture the severity of geographic building restrictions across U.S. housing markets. Unlike previous proxies, the LU-ML indices exhibit strong predictive power for house prices and reveal that looser supply constraints do dampen the house price effects of demand growth, providing long-elusive empirical validation for canonical housing and urban theory.

More specifically, we construct two distinct indices: the LU-ML Price Pressure Index and the LU-ML Supply Index. While both use the same input data and methodology, they capture distinct economic mechanisms. The LU-ML Supply Index targets housing unit growth and thus reveals the relationship between land unavailability and the physical housing stock, akin to conventional supply constraint proxies (Saiz, 2010; Baum-Snow and Han, 2024). Because results using the Supply Index show that looser constraints dampen the price impacts of demand growth, reversing the null findings associated with standard measures (Louie et al., 2025), the Index more accurately

[1] See also The New York Times (2022); Office of the Governor of New York (2023); Office of the Governor of California (2025); Goldman Sachs (2025); United States Senate Committee on Banking, Housing, and Urban Affairs (2025). In the academic literature, see Molloy (2020).

[2] See, for example, Mian and Sufi (2011); Chaney et al. (2012); Mian et al. (2013); Mian and Sufi (2014); Adelino et al. (2015); Aladangady (2017); Chetty et al. (2017); Stroebel and Vavra (2019); Baldauf et al. (2020); Gao et al. (2020); Kaplan et al. (2020a); Griffin et al. (2021); Howard and Liebersohn (2021); Campello et al. (2022); Gupta et al. (2022, 2025); Fonseca et al. (2025).

reflects supply rigidities across housing markets.

However, realized supply is not the only channel through which geography affects markets. Housing is also an asset class, where prices reflect expectations and risk premia (Cheng et al., 2014; Kaplan et al., 2020b; Chopra et al., 2025). Crucially, agents price this risk conditional on local supply constraints. In constrained markets, the inability of supply to absorb shocks increases asset volatility, demanding a higher risk premium (Han, 2013) and altering the option value of development (Titman, 1985; Nathanson and Zwick, 2018).

As the Price Pressure Index is largely unrelated to natural amenity demand—and, by construction, purges other local, idiosyncratic demand and supply shocks—the Index captures this capitalization channel, in addition to the typical supply effects, reflecting the role of geographic supply constraints for house prices. Because prices adjust quickly through changing expectations and reflect persistent risk premia, the mapping from geography to prices differs from the mapping to quantity. By constructing a separate Price Pressure Index, we capture these expectation effects, allowing us to explain heterogeneous price growth over both the short- and long-run.

To build the LU-ML Price Pressure Index, we start with the foundational premise from Saiz (2010) that exogenously unavailable land (due to steep slopes, water, and wetlands) steepens the housing supply curve. Yet unlike Saiz and subsequent work (e.g., Baum-Snow and Han, 2024), which aggregate these three types of LU equally within a single, uniform circle for each city, we measure slope, water, and wetlands separately using multiple spatial definitions both within and around a city. This granularity is crucial: the price impact and risk profile of wetlands differs from that of steep slopes, just as the effects of a constraint at the urban core differ from those at the periphery. Our LU-ML approach captures both of these dynamics and allows for data-driven heterogeneity in how physical features impact local markets.

To compile the LU-ML Price Pressure Index for CBSAs, we use 90 features for each city (slope, water, and wetlands LU measured using 30 different spatial definitions in and around the city). We then train an XGBoost model (Chen and Guestrin, 2016) that uncovers the relationship between physical constraints and prices. The key advantages of XGBoost over standard regression or other linear machine learning methods (e.g., LASSO, PCA) are twofold: (1) it handles the high dimensionality of our 90 distinct features—avoiding the noise and overfitting of standard

regressions, and (2) it captures the rich nonlinearities and interactions that are important for the relationship between LU and house prices. To avoid idiosyncratic local contaminants and isolate the impact of just the LU features, while mitigating model overfitting, we generate the final indices via cross-fitting (Chernozhukov et al., 2018). That is, each city's index value is the out-of-sample (OOS) prediction generated by a model trained exclusively on other markets.

The LU-ML indices offer substantially greater flexibility than existing proxies. They scale to various time periods and multiple levels of disaggregation, from cities (CBSAs) down to zip codes and census tracts, for all housing markets in the contiguous U.S. The LU-ML indices also adapt to various empirical specifications: the model can be trained to target cumulative long-run appreciation for long-difference designs or period-specific price growth for panels. This breadth increases understanding of the relationship between geographic supply constraints and housing markets.

Figure 1 compares the predictive power of the LU-ML Price Pressure Index to the standard Saiz (2010) and Baum-Snow and Han (2024, BSH) elasticity proxies via instrument relevance tests for city-level house prices. Details on the data and specifications are in the Figure notes. Panel A replicates leading studies across various empirical designs, including panels and long differences. In every replication, the BSH proxy yields F-statistics for house prices below 10. The F-statistics for the Saiz proxy clear the traditional threshold of 10 but never reach 50, a modern weak instrument threshold suggested by Keane and Neal (2024). The results are similar for long-differenced specifications across various time periods (Panel B). The Saiz and BSH proxies thus have limited predictive power for house prices, implying that these standard measures of supply constraints are largely unrelated to cross-sectional house price growth. In sharp contrast, the F-statistics for the LU-ML Price Pressure Index exceed the Keane and Neal threshold across all specifications.

Panel C uses our preferred specification:

$$\Delta \log HP_{i,d,t} = \psi_i + \gamma_{d,t} + \beta \cdot SC_{i,d,t} + \varepsilon_{i,d,t} \tag{1}$$

$\Delta \log HP_{i,d,t}$ is the year-over-year (YoY) log first-difference in house prices for CBSA i in Census Division d at month t and SC signifies a supply constraint proxy. Since the dependent variable is in log changes (price growth), time-invariant level factors are differenced away and the CBSA fixed

effects (ψ_i) capture heterogeneous city-level trends. The specification thus accounts for differences in both static city-level attributes, due to amenities or regulation, and secular trends, such as persistent price appreciation in "Superstar Cities" (Gyourko et al., 2013) or the secular boom in the "Sand States" (Davidoff, 2013). The Census division-by-time fixed effects non-parametrically control for shocks to each Census division over time.

$SC_{i,d,t}$ represents the effective, time-varying supply constraint. Because we construct the Price Pressure Index to predict price growth across cities over time, it isolates the variation in local price dynamics driven by the amplification of aggregate shocks in geographically constrained markets. In contrast, the BSH and Saiz proxies are static cross-sectional measures. To make them comparable, we follow the literature and explicitly interact these static proxies with national house price growth (Chaney et al., 2012; Adelino et al., 2015; Aladangady, 2017; Guren et al., 2021; Campello et al., 2022).

Using the panel approach, BSH elasticity is again uncorrelated with house price growth across all sample periods, while Saiz elasticity yields an F-statistic barely above 10 only for the 1976–2024 and the 2000–2011 time periods. Altogether, the results from Figure 1 imply that standard supply constraint proxies are largely uninformative for city-level house price growth, especially after accounting for heterogeneous trends and regional shocks. Conversely, the LU-ML Price Pressure Index has strong predictive power, congruent with geographic supply constraints impacting house prices and validating our LU-ML approach that combines multiple LU estimates via ML techniques. Finally, note that results in Figure 1 extend to house prices measured at various levels of disaggregation (CBSAs, counties, three-digit zip codes, zip codes, and census tracts) and all major publicly available housing datasets.

Davidoff (2016) highlights a key concern when using land unavailability as an exogenous predictor of house prices: households likely find steep-sloped areas and water (e.g., beaches) as valuable natural amenities. Prices in areas with high LU thus might rise not because of construction constraints, but rather due to household demand for these amenities.

The panel structure in Equation 1 addresses Davidoff's LU critique, as the CBSA fixed effects account for city-level trends in amenity demand over time. Nonetheless, we further nonparametrically control for time-varying natural amenity demand—stemming from coastlines, January and

July temperature quintiles, and major parks—by interacting these variables with time fixed effects. We similarly account for natural amenity demand related to hiking, water, and natural recreation, proxied by Google Maps reviews from Lutz (2025b), direct measures of natural amenity demand highlighted by Davidoff. Contrary to the concern that amenity demand drives the predictive power of land unavailability, our results remain robust, and these demand proxies explain little of the variation in the LU-ML Index.

We next investigate the LU-ML Price Pressure Index's prediction errors. Not surprisingly, the LU-ML Index is a poor predictor of house price growth in the tails. Extreme city-level growth often stems from other shocks (e.g., a manufacturing decline or a tech boom). The LU-ML Index, by design, is ill-suited to uncover such shocks. Indeed, we find that city-level total income growth (a standard demand proxy) strongly correlates with the LU-ML residuals, supporting the LU-ML Index as a supply constraint measure rather than a demand proxy.

The strong predictive performance of the LU-ML Price Pressure Index emanates from its integration of diverse LU estimates within a flexible, data-driven framework. A closer examination of the underlying features reveals that steep slopes are the dominant driver of house price growth, far exceeding the impact of water. Economically, this suggests that constraints on the intensive margin—where slopes sharply increase the marginal cost of densification and infrastructure—are more binding for price appreciation than constraints on the extensive margin, such as water boundaries. Wetlands play a negligible role, as they are "soft" constraints often overcome by land-improvement technologies (Dahl, 2011). These results align with Burchfield et al. (2006), who find that topography plays a larger role in shaping development patterns than water or wetlands.

The XGBoost algorithm also uncovers important interactions and nonlinearities that other proxies miss using standard techniques (Kaplan et al., 2020a). We find significant interaction effects between slope and water, as well as a concave relationship between slope ruggedness and prices. Increasing LU due to steep slopes initially raises house prices, but rugged terrain eventually dampens them. This nonlinearity mirrors the results in Burchfield et al. (2006). They find that moderate slopes induce expansive, scattered sprawl, but high mountains act as hard barriers. Our results thus suggest that while moderate topography binds supply (raising prices), extreme slopes eventually limit agglomeration and market access and thus lower the value of the housing stock.

The LU-ML Supply Index follows the same methodology as our Price Pressure Index but uses log change in housing units as the target variable. The LU-ML Supply Index thus isolates the component of housing unit growth driven by supply-side geography. We use the LU-ML Supply Index to address a fundamental housing question: Do looser supply constraints dampen house price growth? Examining existing supply constraint proxies, Louie et al. (2025) conclude that the answer is "No," contrary to the standard theory. They use the following specification across U.S. CBSAs:

$$\begin{aligned}\Delta \log(HP_i) = \alpha + \beta_1 \Delta \log(Income_i) + \beta_2 LessConstrained_i \\ + \beta_3[\Delta \log(Income_i) \times LessConstrained_i] + \epsilon_i\end{aligned} \tag{2}$$

$LessConstrained_i$ is an indicator equal to 1 for CBSAs with looser-than-average supply constraints. For a valid supply constraint proxy, according to standard theory, β_3 must be negative, meaning that looser supply constraints mitigate the house price effects of income shocks. Figure 2 plots the output for separate regressions using the BSH and Saiz proxies, as well as the LU-ML Supply Index. Data and further implementation details are in the Figure notes but generally follow Louie et al. (2025). Panel A first shows the income elasticity of house prices for more constrained CBSAs (β_1) and less constrained CBSAs ($\beta_1 + \beta_3$), while panel B shows the difference in elasticity estimates (β_3). The first two columns of the Figure replicate Louie et al. (2025): using the BSH and Saiz proxies, looser supply constraints do not ameliorate the house price effects of income growth. Furthermore, these proxies imply that less constrained cities do not produce more housing units for a given income shock (Panel C). These results reject the standard narrative that supply elasticity dampens price appreciation.

The third column of Figure 2 shows the results using the LU-ML Supply Index. In marked contrast to other measures, the LU-ML Index reveals a substantial difference in the income elasticity of house prices. A 1% increase in income raises house prices just 0.49% in less constrained cities versus 0.74% in more constrained cities. The difference in elasticities ($\beta_3 = -0.25$) is large in magnitude and statistically significant. Panel C confirms the quantity mechanism using the LU-ML Index: housing units in less constrained cities increase significantly more in response to income growth. These findings extend to multiple house price datasets, to counties, to using the continuous LU-ML Supply Index, and to panel specifications that account for heterogeneous

regional or city-level trends and national business cycles. Such panel specifications yield our preferred estimates, ensuring that elasticity differences are not confounded by secular local supply trends or macroeconomic shocks. Together, these results both validate the LU-ML Indices as supply constraints for housing markets and document that lower construction barriers reduce the price sensitivity of housing markets.

In a novel application, we use the LU-ML Price Pressure Index to revisit the housing collateral channel for entrepreneurship. Exploiting the index's high spatial resolution, we implement a zip-code-level research design that controls for heterogeneous local trends and time-varying state shocks—a level of identification previously unattainable with coarser approaches.[3] We find that rising house prices significantly increase small business formation over the long run (1994–2022). However, this effect evolves over time: it is negligible during the 1990s but strengthens markedly after the Great Recession, suggesting that housing collateral has become increasingly important for entrepreneurs in the modern economy.

1 Data

The U.S. Geological Survey (USGS) provides the two primary datasets necessary to measure land unavailability due to steep slopes, water, and wetlands. The first is the USGS National Elevation Dataset (NED) 3DEP 1 arc-second Digital Elevation Model (DEM). While the original Saiz dataset relied on 3 arc-second DEM data with a resolution of approximately 90 meters, our data provide continuous coverage of the United States at a significantly finer resolution of approximately 30 meters. From the DEM data, we calculate the slope of the terrain to determine the percentage of unavailable land. Following Saiz, we classify any land with a slope exceeding 15% as undevelopable. Our second main dataset is the National Land Cover Database (NLCD) 2011, which uses Landsat imagery to classify land use. From this, we identify categories that constitute unavailable land, specifically water (oceans, lakes, rivers) and wetlands.

Other geospatial data include shapefiles for various census delineations from the NHGIS (Schroeder et al., 2025) and TIGER/Line shapefiles (U.S. Census Bureau, a), and satellite imagery from Google Maps. The National Park Service provides shapefiles for major parks. We obtain Census and weather data from the NHGIS, zip-code-level business patterns from the U.S.

[3] See Chaney et al. (2012), Adelino et al. (2015), Bahaj et al. (2020), and Campello et al. (2022).

Census Bureau (b), and geographic crosswalks from Lutz (2025a). House price datasets are from Freddie Mac, the Federal Housing Finance Agency, and Zillow. The Bureau of Economic Analysis tabulates income at the CBSA and county levels. Finally, journal repositories provide data for replicating Chaney et al. (2012), Mian and Sufi (2014), Stroebel and Vavra (2019), and Guren et al. (2021).

2 The Construction of the LU-ML Indices

The groundbreaking work of Saiz (2010) provides the foundation for this paper, as it was the first to use satellite imagery and GIS methods to compute proxies of LU. Saiz computes the share of land unavailable for construction within a 50km radius around the centroid of each MSA's first central city using 1999 MSA/NECMA definitions. He sums the percentage of land undevelopable due to steep slopes, water, and wetlands to create a single proxy for each city. By linearly summing these features, Saiz effectively assumes that different types of geographic constraints are perfectly substitutable—that water or wetlands impact housing supply identically to steep slopes. This approach is at odds with Burchfield et al. (2006), who find that sloped terrain has a larger impact on development patterns than water or wetlands.

Figure 3, Panel A illustrates the limitations of the Saiz approach using Google satellite imagery for the Los Angeles-Long Beach-Anaheim CBSA. The blue outlined area represents the Los Angeles-Long Beach MSA polygon boundary (2023 delineations) and the red circle marks a 50km radius around the first central city centroid (in this case, the Los Angeles central city). Saiz uses the area within the red circle to calculate land unavailability. Clearly, the location of the first central city centroid determines the geography used in the Saiz calculation: the red circle in panel A captures Los Angeles proper but fails to cover the eastern reaches of the CBSA, as well as northern cities like Lancaster and Palmdale—two large cities with a combined 2000 population of over 230,000—or major economic hubs like Anaheim and Irvine. In the greater LA area, these are the exact regions where much of the new construction occurs.

More generally, the Saiz circle creates systematic, spatially correlated measurement error. It tends to under-cover the spatially expansive CBSAs in the Southwest, while over-covering the more compact jurisdictions in the Northeast. Cities are also heterogeneous along other dimensions, including their polygon shapes, the location of their growth frontiers, and their proximity to

neighboring metros. By applying a uniform approach to diverse markets, the Saiz approach often mischaracterizes the relevant supply constraints facing developers.

To address these limitations, we construct new measurements of LU that scale to various levels of spatial aggregation including CBSAs, counties, three-digit zip codes, zip codes, and census tracts. As differences in geographic boundaries present notable challenges for researchers, an immediate benefit of our dataset is that we match the delineations found in other datasets. This eliminates a vector of uncertainty for researchers studying housing markets.

Noting that any individual land unavailability proxy is likely measured with error, we employ machine-learning techniques to combine land unavailability estimates measured in and around each geographic boundary.

Panels B–D in Figure 3 show an example of our approach for the Los Angeles-Long Beach-Anaheim CBSA. We consider various buffered polygons around the first principal city (Panel B), buffers around the CBSA polygon (Panel C), and circles around the principal city centroid (Panel D).

In Panel B, we calculate land unavailability within the first principal city polygon (by population size; orange polygon), corresponding to the Los Angeles principal city. Then, we buffer this polygon by 10% (inner yellow polygon) and calculate land unavailability. From there, we sequentially increase the buffer size by 10pp until the buffer reaches 150% (outer yellow polygon) of the original first principal city polygon. The result is 16 polygons around the first principal city. These polygons allow us to directly capture land unavailability at the population-weighted center of a CBSA.

We apply the same approach to the overall CBSA polygon (Figure 3, Panel C), yielding five additional polygons corresponding to the CBSA polygon with buffers ranging from 0% to 20% (in 5pp increments). The advantage of buffering the CBSA polygon is that it covers the entire administrative definition of the market and accounts for idiosyncrasies in the polygon's shape. Finally, in Panel D, we expand Saiz's original approach by considering multiple circles around the first principal city centroid instead of just within a 50km radius circle: we exploit nine distinct circles with radii around the first principal city centroid ranging from 20 to 100km (by 10km increments). In the case of Los Angeles in Panel D, larger circles encompass both northern and

southern LA, including the cities of Anaheim, Irvine, Lancaster, and Palmdale missed by the Saiz circle. Overall, Figure 3 exemplifies how we calculate land unavailability for 30 different polygons for each CBSA. However, the best land unavailability predictor for house prices is likely a combination of individual land unavailability estimates.

The second step of our approach synthesizes these disaggregated land unavailability (LU) estimates using machine learning (ML) techniques. Specifically, we construct two complementary metrics: the LU-ML Price Pressure Index and the LU-ML Supply Index. While both LU-ML indices exploit the same methodology and LU estimates—steep slopes, water, and wetlands, measured at various levels of disaggregation—we build them from different outcome variables. The LU-ML Price Pressure Index targets house price growth, isolating how physical supply constraints push up prices. The LU-ML Supply Index instead targets the log change in housing units, capturing the direct relationship between geographic barriers and the physical housing stock.

The LU-ML framework also accommodates various empirical settings. Depending on the econometric specification, the algorithm generates either (1) a time-invariant cross-sectional proxy for each geographic unit; or (2) a time-varying panel of LU-based predictors.

Here, we describe our methodology within a panel data setup for the LU-ML Price Pressure Index targeting CBSA YoY house price growth, but this framework easily extends to long differences, other spatial definitions, and the LU-ML Supply Index.

Because the economic impact of physical constraints may vary over time (e.g., constraints may bind more tightly during booms), we estimate the model separately for each time period. We use a repeated cross-fitting approach (Chernozhukov et al., 2018). For each time period in the panel, we implement 5 repeats of 5-fold cross-fitting. Within each fold, we train the XGBoost algorithm on a training set and predict house prices out-of-sample for the holdout (test) set. The average prediction across all repeats for a given CBSA and time period constitutes the final index value. The XGBoost algorithm handles many correlated or weak predictors, unlike standard regression techniques, allowing the data to determine the functional form for the relationship between geographic supply constraints and prices. XGBoost also allows the LU-ML indices to capture rich interactions and nonlinear effects that linear machine learning specifications (such as LASSO or PCA) would fail to detect.

Specifically, for each time period t in the panel dataset:

1. Randomly split the data into 5 equal folds.
2. Designate one fold as the holdout (test) set, where the remaining folds constitute the training set.
3. Train the XGBoost algorithm on the training data.[4]
4. Predict house price growth out-of-sample for the holdout set.
5. Repeat steps (2) to (4), rotating which fold serves as the holdout set.
6. Repeat steps (1) to (5) 5 times (5 repeats) with different random splits.
7. Average the out-of-sample house price growth predictions across all repeats by CBSA to generate the LU-ML Price Pressure Index for time period t.

Repeating this process for every period t yields a panel LU-ML Index. If the target variable is long-differenced, we generate a cross-sectional LU-ML Index by following steps (1) to (7) once. Using out-of-sample predictions via cross-fitting mitigates overfitting and also prevents local, idiosyncratic supply and demand factors from contaminating the LU-ML Indices.

3 The LU-ML Price Pressure Index and House Prices

Figure 4 summarizes the predictive power of the LU-ML Price Pressure Index relative to the BSH and Saiz supply constraint proxies using instrument relevance tests for house price growth. Specifically, we expand on the panel CBSA-level results in Figure 1—which used Freddie Mac data—by aggregating results across various housing datasets (FHFA, Freddie Mac, and Zillow) at several levels of disaggregation (CBSA, county, three-digit zip code, and census tract) using Equation 1. Panel A reports the median first-stage F-statistic by estimation period, Panel B plots the median partial R^2, and Panel C displays the share of cross-sectional units covered by each proxy.

Panels A and B show that existing supply constraint proxies have limited predictive power for house prices. The BSH measure has negligible F-statistics across all estimation periods. The Saiz proxy has marginally better predictive performance but consistently falls short of the $F = 50$ weak instrument threshold (black-dashed line) suggested by Keane and Neal (2024) and often fails to

[4]To prevent overfitting during training, we further partition the training set to create a validation set. We use this set to tune the learning rate and implement early stopping (XGBoost Documentation, 2025), which halts the training process if the prediction error fails to improve after a specified number of rounds. We use default values for all other hyperparameters.

clear the traditional $F = 10$ threshold. Consequently, these standard proxies explain little of the cross-sectional variation in house price growth (Panel B).

Unlike the BSH and Saiz proxies, the LU-ML Price Pressure Index exhibits robust predictive power, with large F-statistics and partial R^2 values across specifications. These results thus document that geographic supply constraints do explain variation in the cross-section of house price growth, congruent with the canonical view that supply constraints push up house prices.

Panel C highlights another advantage of the LU-ML Index: its geographic coverage. The BSH and Saiz proxies often exclude 20–40% of available cross-sectional units, with samples skewed toward larger, more established metros. Empirical estimates using these proxies may thus suffer from sample selection bias. The LU-ML Indices, conversely, have nearly 100% coverage, only discarding Alaska and Hawaii.

Appendix Tables A8 and A9 confirm that the LU-ML Price Pressure Index outperforms existing proxies across every house price dataset and spatial aggregation tested. Finally, Appendix Figure A3 documents that these results extend to long-differenced specifications.

3.1 The LU-ML Index and Natural Amenity Demand

Davidoff (2016) challenges the validity of geographic supply proxies, arguing that the positive correlation between land unavailability and prices reflects household demand for natural amenities (e.g., water access, views) rather than supply constraints. Using a multifaceted approach, we find that the LU-ML Indices reflect supply constraints rather than amenity demand.

First, our preferred specification (Equation 1) addresses the core of the Davidoff critique. As the dependent variable is price growth ($\Delta \log HP$), time-invariant amenity values are differenced out. Furthermore, the inclusion of CBSA fixed effects (ψ_i) absorbs heterogeneous city-level trends, including secular shifts in the value of local amenities.

Table 1, Column (1) reports estimates from Equation 1 using the Freddie Mac, Zillow, and FHFA house price datasets, controlling only for CBSA and Census division $\times$ time fixed effects. Across all datasets, the coefficient on the LU-ML Price Pressure Index is large and statistically significant, consistent with LU-induced supply constraints driving house price growth beyond both static and trending amenity demand.

Columns (2)–(4) add traditional proxies for natural amenity demand highlighted by Davidoff

and others. We include a coastal indicator, mean January and July temperature quintiles, and a major park indicator using data from the National Park Service, interacting these proxies with time fixed effects. This flexible specification allows the capitalization of these amenities into house prices to evolve nonparametrically over time. While Davidoff's critique implies that the predictive power of the LU-ML Index should vanish in the presence of these controls, the estimated effect of the Price Pressure Index on house price growth instead persists and remains highly significant.

Column (5) directly accounts for natural amenities by controlling for the "Google Amenity Demand Indices" from Lutz (2025b). These indices are derived from millions of Google Maps reviews for hiking, water access, and natural recreation and thus capture revealed preferences for natural amenities. Even after controlling for these demand proxies interacted with time fixed effects, the relationship between the LU-ML Price Pressure Index and house price growth remains robust.

Finally, we find that the Google Amenity Demand Indices explain only 2–5% of the variation in the LU-ML Price Pressure Index from 2002–2024 (Adjusted R^2), further countering Davidoff's contention that measures of land unavailability are merely proxies for amenity demand.[5] Altogether, these results suggest that LU-ML Indices measure the effects of supply constraints rather than amenity demand.

3.2 The LU-ML Price Pressure Index Prediction Errors

As a supply constraint measure, the LU-ML Price Pressure Index should capture the cross-sectional price variation driven by physical building limitations while remaining largely unrelated to local demand shocks. By design, the use of cross-fitting insulates the LU-ML Index from such shocks. We therefore expect the Index to exhibit systematic prediction errors in the tails of the house price distribution, where large demand shocks—such as a tech boom or a manufacturing collapse—often dominate local price dynamics.

Figure 5, Panel A confirms this pattern by plotting the LU-ML OOS prediction errors across CBSA house price growth bins from 2002–2024. We calculate errors as the actual Zillow CBSA house price growth minus the LU-ML prediction. In the middle of the distribution, where demand

[5]Results stem from a long-difference regression of the LU-ML Price Pressure Index on the Google Amenity Demand Indices, both separately and together. To compile the long-differenced LU-ML Price Pressure Index, we use the long difference in Zillow house price growth from 2002–2024 as the target variable in our LU-ML approach.

shocks are often moderate, the LU-ML prediction errors are small. The tails, however, yield sizable residuals. In the highest growth bin (far right), the prediction errors are large and positive: actual appreciation far exceeded what geographic constraints alone would predict, consistent with an outsized positive demand shock. Conversely, in the lowest growth bin, errors are large and negative, congruent with a negative demand shock pulling prices below their geography-implied levels. Overall, the monotonic relationship in Panel A indicates that local demand shocks drive the LU-ML prediction errors.

Panel B explicitly examines the LU-ML residuals and housing demand. Here, we plot a binscatter of the LU-ML Index's OOS prediction errors against CBSA total income growth (a standard demand proxy) from 2002–2023. The relationship is positive and highly significant ($t = 3.14$), confirming that demand captures the price variation missed by the LU-ML Index. This further supports the LU-ML Index as a supply-side constraint measure that leaves demand shocks in the residual.

3.3 Decomposing the Geographic Drivers of Price Pressure

Having established that geographic constraints are important determinants of house price growth—contrary to inferences drawn from prior proxies—we next open the "black box" to uncover the economic mechanisms driving these results. Specifically, we investigate which physical constraints amplify prices. For example, steep slopes bind on the intensive margin by raising the marginal cost of density, whereas water boundaries constrain growth along the extensive margin as an urban border (Burchfield et al., 2006). Distinguishing between these mechanisms clarifies whether high prices stem from physical limits on a city's footprint or from the escalating marginal costs of densification.

We decompose the city-level LU-ML Price Pressure Index into its constituent geographic drivers using SHAP (SHapley Additive exPlanations) values (Shapley, 1953; Lundberg and Lee, 2017; Lundberg et al., 2020). This approach isolates the marginal contribution of each LU feature—steep slopes, water, and wetlands across various spatial definitions—to the final LU-ML Index. The SHAP value thus represents the specific growth premium (or discount) attributable to a single geographic feature. As derived by Shapley (1953) in the context of cooperative game theory, SHAP values are additive: the LU-ML Index for each city is the sum of the national average plus the

marginal contributions of its specific geographic components.

Formally, let $\varphi_i^{(j)}$ denote the SHAP value for LU feature j in CBSA i: the average marginal impact of feature j relative to all possible coalitions of other features for city i. Here, j may represent a main effect or an interaction term. We can decompose the LU-ML Price Pressure Index, the model's total predicted price growth for city i, as:

$$\text{LU-ML Index}_i = \mu + \sum_{j=1}^{M} \varphi_i^{(j)} \tag{3}$$

where μ represents the unconditional mean prediction across all CBSAs. A positive SHAP value ($\varphi_i^{(j)} > 0$) implies that LU feature j acts as a binding supply constraint for CBSA i, pushing its Price Pressure Index above the national baseline. Conversely, a negative SHAP value signals a looser-than-average supply constraint for feature j. This decomposition thus ranks geographic constraints by their economic magnitude, city-by-city.

Figure 6 decomposes the LU-ML Price Pressure Index—based on 2002–2024 Zillow house price growth—for representative cities. For each metro, we report the LU-ML Index, the excess pressure relative to the national baseline ($\sum \varphi_i^{(j)}$), and the prediction residual.[6] The arrows in the plot display aggregated SHAP values: the total marginal contribution of each LU type—steep slopes (red), water (blue), and wetlands (teal)—to the LU-ML Price Pressure Index.[7] The arrow labels report the share of undevelopable land (averaged across spatial definitions) and the corresponding national percentile by LU type.

The figure highlights the substantial heterogeneity in how physical constraints manifest across different markets. Houston has the lowest Price Pressure Index (99%), consistent with a geographic landscape that offers little resistance to supply. Its lack of steep-sloped terrain pulls price growth 22pp below the national average. In contrast, Salt Lake City has the highest Index (183%), where both rugged terrain (+31pp) and water barriers (+18pp) bind severely. San Francisco (LU-ML Index = 165%) and Miami (LU-ML Index = 146%) also exhibit high geographic pressure, though the specific drivers differ.

The SHAP values further reveal that geographic constraints are not monolithic; their impact

[6] Because predictions are generated out-of-sample, $\hat{\mu}$ varies slightly across estimation folds, ranging from 130–136% with a mean of 134%.

[7] We calculate the aggregate impact for each broad category by summing the SHAP values by LU type (e.g., $\text{Slope Impact}_i = \sum_{j \in \text{Slope}} \varphi_i^{(j)}$).

is nonlinear and context-dependent. We observe three stylized facts. First, steep slopes are the dominant constraint for many markets. In flat cities like Houston, Miami, and Oklahoma City, the absence of slopes represents supply slack, pulling the Price Pressure Index 10–22pp below the national average. Conversely, in Salt Lake City, rugged terrain (slope: 36%, 88th percentile) boosts predicted growth by over 30pp.

Second, the impact of slope on prices appears non-monotonic. San Francisco possesses significantly more LU due to slopes (48%) than Salt Lake City (36%), but the marginal contribution of slope to price growth is smaller (+9.7pp vs +31.1pp). This suggests an inverted U-relationship: while moderate constraints raise the marginal cost of density (increasing prices), extreme constraints may limit development so severely that the marginal pressure on prices diminishes. We examine this nonlinearity further below in Figure 7.

Third, water constraints appear to bind only at extremes. Consistent with Burchfield et al. (2006), who find that water boundaries generally do not constrain development, Figure 6 documents that water exerts downward price pressure in Houston (−8.7pp) despite its relatively high water coverage (6%, 82nd percentile). However, our results also show that this relationship reverses at the tail. When water features reach extreme levels that often define hard urban boundaries—as in Miami, San Francisco, and Salt Lake City—they become binding supply constraints, adding 18–29pp to the LU-ML Index.

Finally, as we construct the LU-ML Indices using out-of-sample cross-fitting, the residuals (Actual Growth − LU-ML Index) largely reflect demand-side shocks rather than geographic supply constraints (Figure 5). Consequently, the large positive residuals in Miami (+70pp) and Houston (+37pp) stem from idiosyncratic demand shocks—such as COVID-era migration booms (Hudson, 2025) or recent affordability crises (Glaeser and Gyourko, 2025; Sherman et al., 2025)—beyond the price growth predicted by terrain alone.

3.4 The Structure of Price Pressure Across Cities

Figure 7 decomposes the LU-ML Price Pressure Index to identify the specific features, functional forms, and spatial scales that drive geographic price pressures. Panel 1 plots the marginal contribution (aggregated SHAP values) against the LU share for each feature type. These plots confirm that the stylized facts discussed above extend to the full distribution of cities.

Panel 1A documents that the relationship between topography and price pressure is nonlinear. In flat cities, the absence of slope exerts downward pressure on the LU-ML Index (20pp below the national mean), consistent with highly elastic housing supply. As terrain becomes rugged, the price pressure rises sharply, consistent with slopes acting as a binding constraint on the intensive margin (Burchfield et al., 2006). However, the marginal contribution of steep slopes is concave: beyond an LU share of approximately 40%, the marginal impact diminishes and eventually reverses. This shape suggests that while moderate ruggedness constrains supply and raises prices, extreme topography may eventually sever market access or reduce agglomeration benefits, thereby dampening valuations.

Water, by contrast, exhibits a different pattern (Panel 1B). At low levels, the marginal contributions of water LU are widely dispersed, congruent with Burchfield et al. (2006) who find that water often does not impact development patterns. However, high levels of water push prices upwards. This finding is consistent with water acting as a hard boundary to extensive margin growth. Panel 1C confirms that wetlands, conversely, act as a "soft constraint" with negligible price impacts (Dahl, 2011).

Panel 2 ranks the drivers of geographic price pressure by economic magnitude. In Panel 2A, we group features by LU type but also allow for interactions. This plot confirms that steep slopes are the dominant constraint, with an absolute impact of 20pp on the LU-ML Index for the average city relative to the national baseline. The second most important LU type is water (10pp). Interaction effects are also economically meaningful. The slope $\times$ water interaction (6.8pp) and the slope $\times$ wetlands (6pp) together exert more price pressure than water alone. These results empirically validate the use of XGBoost: a significant portion of geographic price pressure stems from the confluence of constraints, rather than just their additive sum.

Finally, Panel 2B identifies the relevant spatial scale of these constraints. We find that the 100km radius around the principal city centroid is the single most important spatial predictor (16.9pp). The dominance of this broad measure—relative to the standard 50km radius used in Saiz (2010)—suggests that modern housing supply constraints are determined at the regional level rather than solely by local bottlenecks. However, the non-trivial contributions of smaller buffers highlight that constraints bind simultaneously at the urban core and the periphery.

In sum, the LU-ML Price Pressure Index demonstrates that geographic supply constraints are neither linear nor strictly local. They operate through concave nonlinearities and threshold effects, compound via interactions, and bind most tightly at the regional periphery.

4 The LU-ML Supply Index and the Income Elasticity of House Prices

We next test whether looser supply constraints mitigate the house price effects of demand shocks. To do so, we employ the LU-ML Supply Index. As described in Section 2, this index shares the same high-resolution LU inputs and ML methodology as the Price Pressure Index but targets the log change in housing units. It therefore sorts housing markets based on the responsiveness of the physical housing stock to geographic constraints.

Using Equation 2 and the LU-ML Supply Index, Figure 2 documents that looser geographic supply constraints mitigate the house price impacts of demand shocks. Specifically, the income elasticity of house prices is roughly 50% higher in constrained markets compared to their unconstrained counterparts (0.74 vs 0.49). This finding resolves the disconnect in the recent literature, where estimates using standard proxies (e.g., BSH and Saiz) imply that more lax supply constraints do not dampen the price appreciation induced by income shocks (Louie et al., 2025). In this section, we demonstrate that our results using the LU-ML Supply Index are a general feature of U.S. housing markets: the mitigating effect of looser supply constraints holds across alternative house price datasets, the full distribution of U.S. cities and counties, alternative sample periods, and various panel specifications that control for local secular trends.

Figure 8 shows the elasticity estimates from equation 2 across various house price datasets and geographies. Panel 1 benchmarks elasticity estimates using the LU-ML Supply Index against the BSH and Saiz proxies across three house price datasets (Freddie Mac, FHFA, and Zillow) for the 2000–2020 long-difference period. To ensure a consistent comparison, we restrict the sample to major metros. Panel 1A plots the elasticities for constrained cities, while Panel 1B isolates the differential impact of supply constraints (β_3).

Consistent with Louie et al. (2025), the elasticity differentials for BSH (green) and Saiz (red) in Panel 1B are small and not statistically significant. This confirms that existing proxies fail to capture the supply-side mitigation effects of demand shocks. In sharp contrast, the LU-ML Supply Index (blue) yields a large, statistically significant, and negative coefficient that is consistent across

all three datasets, highlighting the robustness of the findings using the LU-ML Index.

A distinct advantage of the LU-ML approach is its scalability. While standard proxies are largely confined to major metropolitan areas, our index covers the contiguous U.S. Panel 2 exploits this coverage to estimate the elasticity differentials using the comprehensive FHFA dataset with nearly all U.S. CBSAs and counties. Even when expanding the geographic sample or aggregating data to counties, the interaction coefficient (β_3) remains negative and statistically significant (Panel 2B). This confirms that geography-induced supply constraints shape price dynamics not just in "Superstar Cities," but across the broader U.S. housing market.

Panel 3 implements our preferred specification using a decadal panel from 1990-2020:

$$\Delta \log(HP_{it}) = \alpha_i + \eta_t + \beta_1 \Delta \log(Income_{it}) + \beta_3[\Delta \log(Income_{it}) \times LessConstrained_i] + \epsilon_{it} \quad (4)$$

where α_i and η_t signify location (CBSA or county) and decadal-time fixed effects. Because the dependent variable is the log change in prices, the location fixed effects control for heterogeneous city- or county-specific growth, such as local supply trends or the secular appreciation in "Superstar Cities" and the "Sand States". This specification thus exploits within-location variation to test if supply constraints mitigate the price effects of demand shocks relative to a city-level baseline, further isolating the supply constraint mechanism. Estimates of β_3 remain negative, economically large, highly significant, and consistent across both CBSAs and counties.

The previous estimates relied on a binary classification of constrained versus unconstrained markets. In Table 2, we exploit the granularity of our measure within a 1990–2020 decadal panel by interacting income growth with the continuous LU-ML Supply Index (standardized to have zero mean and unit variance). Column 1 controls only for time fixed effects, mimicking a stacked cross-sectional design. Columns 2–4 progressively add state, commuting zone, and CBSA fixed effects.

Two key patterns emerge. First, the estimates for β_3 are consistently negative and significant: a one standard deviation increase in the LU-ML index (i.e., looser constraints) reduces the income elasticity of house prices by 0.09 to 0.15. Second, the magnitude of this effect increases as we control for finer geographic fixed effects. This implies that unobserved local heterogeneity tends to bias the supply constraint effect toward zero. Controlling for these local factors hence reveals that geography plays an even larger role in dampening price shocks than the national comparisons

used in previous literature suggest.

We also note that the coefficient on the Supply Index level term is positive (0.046 in Column 1). This indicates that while elastic markets are less sensitive to income shocks, they experienced higher average price growth over this period. This result is consistent with Davidoff (2013) and Glaeser and Gyourko (2025), who find that traditionally elastic markets, such as the "Sand States," have experienced significant secular booms. Indeed, Appendix Tables A12 and A13 confirm that non-coastal cities in the Sand States and Texas experienced rapid appreciation despite looser constraints.

Finally, we relax the assumption that supply constraints are time-invariant—a key innovation over static existing proxies. In Table 3, we construct a time-varying panel LU-ML Supply Index, targeting housing unit growth separately for each decade. The results confirm our core findings: looser geographic constraints significantly dampen the house price effects of demand shocks. These estimates, robust across aggregation levels and time periods, provide strong empirical validation of our LU-ML approach.

5 Application: The Housing Collateral Channel and Entrepreneurship

In a novel application, we use the LU-ML Price Pressure Index to revisit a central question in macro-finance: Does rising housing wealth fuel entrepreneurship? Theory suggests that higher collateral values relax borrowing constraints, enabling small business formation (Chaney et al., 2012; Adelino et al., 2015). However, causally identifying this channel is challenging. Omitted local factors like gentrification or agglomeration shocks can simultaneously drive house prices and firm entry. Furthermore, firm formation itself can raise local prices (reverse causality), while measurement error in price indices biases naive OLS estimates.

We overcome this identification challenge by leveraging the high spatial resolution of the LU-ML Price Pressure Index. Unlike coarse county-level instruments, our zip-code-level proxy allows us to control for granular fixed effects and trends. We estimate the following 2SLS specification using zip-code-level data from the Census Business Patterns for 1994–2022:

$$\Delta \log(\text{Firms}_{i,s,t}^{1-4}) = \alpha_i + \delta_{s,t} + \beta \cdot \Delta \log(\text{HP}_{i,s,t}) + \epsilon_{i,s,t} \tag{5}$$

where $\Delta \log(\text{Firms}_{i,s,t}^{1-4})$ is the year-over-year log change in the number of firms with 1–4 employees

in zip code i and state s at year t. Crucially, the inclusion of zip code fixed effects (α_i) in a growth equation absorbs heterogeneous local trends in firm formation, while state-year fixed effects ($\delta_{s,t}$) non-parametrically control for time-varying regional shocks. We instrument house price growth ($\Delta \log(\text{HP}_{i,s,t})$) using the LU-ML Price Pressure Index. The first-stage F-statistic is 614, confirming that local supply constraints are powerful predictors of prices even at this high level of disaggregation.

Figure 9 presents the 2SLS results. Panel A documents a strong causal link between collateral values and micro-entrepreneurship. We find a house price elasticity of 0.075 for firms with 1–4 employees, significant at the 1% level. To validate that this effect is driven by credit constraints rather than general local demand, we run a placebo test using slightly larger firms (5–9 employees). Consistent with the collateral channel hypothesis—which predicts that home equity matters most for the smallest, most capital-constrained entrants—the elasticity for 5–9 employee firms is statistically indistinguishable from zero. The difference between the two estimates is large and statistically significant, confirming that the housing wealth effect is concentrated at the extensive margin of small firm formation.

Panel B investigates the stability of this relationship over the last three decades. We find that the elasticity of entrepreneurship has evolved significantly. The effect was statistically insignificant during the 1990s and modest during the pre-GFC boom (2001–2012). However, the elasticity effectively triples in the post-crisis period (2013–2019) and remains elevated through the pandemic era (2020–2022). This structural shift suggests that housing collateral has become an increasingly binding constraint for entrepreneurs in the modern economy, potentially reflecting tighter small business lending standards or the changing nature of self-employment.

6 Conclusion

This paper revisits the fundamental relationship between physical geography and housing markets. By combining high-resolution satellite imagery with machine learning, we construct the LU-ML Indices—a new class of supply constraint measures that capture the heterogeneity, non-linearities, and interactions of physical barriers at a granular level. In doing so, we resolve a growing disconnect between urban theory and empirical evidence, providing robust validation for the role of supply constraints in shaping U.S. housing markets.

Our analysis yields three primary insights. First, we find that the LU-ML Price Pressure Index predicts house price growth across cities, unlike existing proxies, establishing that supply constraints are a key determinant of cross-sectional price growth. Decomposing the Index reveals that constraints on the intensive margin—specifically steep slopes that raise the marginal cost of density—are the dominant driver of price pressure, far outweighing the impact of water boundaries. Furthermore, we uncover critical non-linearities: while moderate ruggedness steepens the supply curve, extreme topography eventually hampers price growth, creating a non-monotonic relationship that linear models fail to capture.

Second, we document the importance of supply constraints in empirical housing research. Louie et al. (2025) use standard proxies and find that looser supply constraints do not mitigate the house price effects of income shocks. We show that this "null result" is likely an artifact of measurement. Using the LU-ML Supply Index, we document a statistically significant and economically large divergence in elasticities: income shocks trigger substantially smaller price increases in less constrained markets. This finding confirms the canonical theoretical prediction that physical barriers to construction are active amplifiers of prices.

Third, the granular resolution of the LU-ML approach unlocks new identification strategies. In a novel application to entrepreneurship, we exploit the Index's zip-code-level variation to control for local trends—a level of identification unattainable with coarser approaches. We find that the elasticity of small firm formation to housing wealth has increased markedly since the Great Recession, suggesting an increased reliance on home equity for small business formation.

These findings have important implications for policy and future research. By accurately quantifying the impacts of physical constraints, the LU-ML Indices allow for a more precise isolation of other effects, such as regulation. As policymakers target land use reform, distinguishing between the immutable constraints of geography and the malleable constraints of regulation will be essential for designing effective interventions.

References

M. Adelino, A. Schoar, and F. Severino. House prices, collateral, and self-employment. *Journal of Financial Economics*, 117(2):288–306, 2015.

A. Aladangady. Housing wealth and consumption: Evidence from geographically linked microdata. *American Economic Review*, 107(11):3415–3446, 2017.

S. Bahaj, A. Foulis, and G. Pinter. Home values and firm behavior. *American Economic Review*, 110(7):2225–2270, 2020.

M. Baldauf, L. Garlappi, and C. Yannelis. Does climate change affect real estate prices? only if you believe in it. *The Review of Financial Studies*, 33(3):1256–1295, 2020.

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

M. Burchfield, H. G. Overman, D. Puga, and M. A. Turner. Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics*, 121(2):587–633, 2006.

Bureau of Economic Analysis. Personal Income. URL https://www.bea.gov/data/income-saving/personal-income.

M. Campello, R. A. Connolly, G. Kankanhalli, and E. Steiner. Do real estate values boost corporate borrowing? evidence from contract-level data. *Journal of Financial Economics*, 144(2):611–644, 2022.

M. D. Cattaneo, R. K. Crump, M. H. Farrell, and Y. Feng. On binscatter. *American Economic Review*, 114(5):1488–1514, 2024.

T. Chaney, D. Sraer, and D. Thesmar. The collateral channel: How real estate shocks affect corporate investment. *American Economic Review*, 102(6):2381–2409, 2012.

T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785–794, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450342322. doi: 10.1145/2939672.2939785. URL https://doi.org/10.1145/2939672.2939785.

I.-H. Cheng, S. Raina, and W. Xiong. Wall street and the housing bubble. *American Economic Review*, 104(9):2797–2829, 2014.

V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 01 2018. ISSN 1368-4221. doi: 10.1111/ectj.12097. URL https://doi.org/10.1111/ectj.12097.

R. Chetty, L. Sándor, and A. Szeidl. The effect of housing on portfolio choice. *The Journal of Finance*, 72(3):1171–1212, 2017.

F. Chopra, C. Roth, and J. Wohlfart. Home price expectations and spending: Evidence from a field experiment. *American Economic Review*, 115(7):2267–2305, 2025.

T. E. Dahl. *Status and trends of wetlands in the conterminous United States 2004 to 2009*. U.S. Department of the Interior; Fish and Wildlife Service, Washington, D.C., 2011.

T. Davidoff. Supply elasticity and the housing cycle of the 2000s. *Real Estate Economics*, 41(4): 793–813, 2013.

T. Davidoff. Supply constraints are not valid instrumental variables for home prices because they are correlated with many demand factors. *Critical Finance Review*, 5(2):177–206, 2016.

Federal Housing Finance Agency. FHFA House Price Index (HPI). URL https://www.fhfa.gov/data/hpi.

J. Fonseca, L. Liu, and P. Mabille. Unlocking mortgage lock-in: Evidence from a spatial housing ladder model. *Working Paper*, 2025.

Freddie Mac. Freddie Mac House Price Index (FMHPI). URL https://www.freddiemac.com/research/indices/house-price-index.

Z. Gao, M. Sockin, and W. Xiong. Learning about the neighborhood. *The Review of Financial Studies*, 34(9):4323–4372, 11 2020. ISSN 0893-9454. doi: 10.1093/rfs/hhaa130. URL https://doi.org/10.1093/rfs/hhaa130.

E. L. Glaeser and J. Gyourko. America's housing supply problem: The closing of the suburban frontier? Technical report, National Bureau of Economic Research, 2025.

Goldman Sachs. The outlook for US housing supply and affordability, Oct. 2025. URL https://www.goldmansachs.com/insights/articles/the-outlook-for-us-housing-supply-and-affordability. Macroeconomics Insights.

Google Maps. Google Maps Platform. URL https://developers.google.com/maps.

J. M. Griffin, S. Kruger, and G. Maturana. What drove the 2003–2006 house price boom and subsequent collapse? disentangling competing explanations. *Journal of Financial Economics*, 141(3):1007–1035, 2021.

A. Gupta, V. Mittal, J. Peeters, and S. Van Nieuwerburgh. Flattening the curve: pandemic-induced revaluation of urban real estate. *Journal of Financial Economics*, 146(2):594–636, 2022.

A. Gupta, H. Sapriza, and V. Yankov. The collateral channel and bank credit. *Working Paper*, 2025.

A. M. Guren, A. McKay, E. Nakamura, and J. Steinsson. Housing wealth effects: The long view. *The Review of Economic Studies*, 88(2):669–707, 2021.

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

L. Han. Understanding the puzzling risk-return relationship for housing. *The Review of Financial Studies*, 26(4):877–928, 2013.

G. Howard and J. Liebersohn. Why is the rent so darn high? the role of growing demand to live in housing-supply-inelastic cities. *Journal of Urban Economics*, 124:103369, 2021.

T. Hudson. 'The risk levels are higher': Miami is the number one housing market bubble in the world – again. *The Miami Times*, Oct. 2025.

G. Kaplan, K. Mitman, and G. L. Violante. Non-durable consumption and housing net worth in the great recession: Evidence from easily accessible data. *Journal of Public Economics*, 189: 104176, 2020a.

G. Kaplan, K. Mitman, and G. L. Violante. The housing boom and bust: Model meets evidence. *Journal of Political Economy*, 128(9):3285–3345, 2020b.

M. P. Keane and T. Neal. A practical guide to weak instruments. *Annual Review of Economics*, 16, 2024.

S. Louie, J. A. Mondragon, and J. Wieland. Supply constraints do not explain house price and quantity growth across us cities. *Working Paper*, 2025.

S. M. Lundberg and S.-I. Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.

S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee. From local explanations to global understanding with explainable ai for trees. *Nature machine intelligence*, 2(1):56–67, 2020.

C. Lutz. Bridging the geographic divide: Crosswalks across space and time. *Working Paper*, 2025a.

C. Lutz. Natural amenity demand index construction using google places. https://github.com/ChandlerLutz/gmaps-natural-amenities-2025, 2025b.

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

A. Mian and A. Sufi. What explains the 2007–2009 drop in employment? *Econometrica*, 82(6): 2197–2223, 2014.

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

R. Molloy. The effect of housing supply regulation on housing affordability: A review. *Regional science and urban economics*, 80(C):1–5, 2020.

C. G. Nathanson and E. Zwick. Arrested development: Theory and evidence of supply-side speculation in the housing market. *The Journal of Finance*, 73(6):2587–2633, 2018.

National Park Service. GIS and Mapping: Tools and Data. URL https://www.nps.gov/subjects/gisandmapping/tools-and-data.htm.

Office of the Governor of California. Governor newsom signs into law groundbreaking reforms to build more housing, boost affordability, June 2025. URL https://www.gov.ca.gov/2025/06/30/governor-newsom-signs-into-law-groundbreaking-reforms-to-build-more-housing-affordabilit Press Release.

Office of the Governor of New York. The new york housing compact, Mar. 2023. URL https://www.governor.ny.gov/sites/default/files/2023-03/NY_Housing_Compact_info_deck_March_2023.pdf. Information Deck.

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

J. Schroeder, D. Van Riper, S. Manson, K. Knowles, T. Kugler, F. Roberts, and S. Ruggles. IPUMS National Historical Geographic Information System: Version 20.0 [dataset], 2025. URL http://doi.org/10.18128/D050.V20.0.

L. S. Shapley. A value for n-person games. 1953.

S. A. Sherman, C. S. Cheong, D. Banerjee, A. Kim, and A. Yang. The 2025 State of Housing in Harris County and Houston. Report, Kinder Institute for Urban Research, Rice University, 2025. URL https://kinder.rice.edu/research/2025-state-housing-harris-county-and-houston.

J. H. Stock and M. Yogo. Testing for weak instruments in linear iv regression. In D. W. K. Andrews and J. H. Stock, editors, *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, pages 80–108. Cambridge University Press, New York, 2005.

J. Stroebel and J. Vavra. House prices, local demand, and retail prices. *Journal of Political Economy*, 127(3):1391–1436, 2019.

The New York Times. The housing shortage isn't just a coastal crisis anymore. *The New York Times*, July 2022. URL https://www.nytimes.com/2022/07/14/upshot/housing-shortage-us.html. The Upshot.

The White House. Fact sheet: President donald j. trump delivers emergency price relief for american families to defeat the cost-of-living crisis, Jan. 2025. URL https://www.whitehouse.gov/fact-sheets/2025/01/fact-sheet-president-donald-j-trump-delivers-emergency-price-relief-for-american-familie Fact Sheet.

S. Titman. Urban land prices under uncertainty. *The American Economic Review*, 75(3):505–514, 1985.

United States Senate Committee on Banking, Housing, and Urban Affairs. ROAD to housing act of 2025: Section-by-section, July 2025. URL https://www.banking.senate.gov/imo/media/doc/road_to_housing_act_of_2025_section_by_section.pdf. Committee Document.

U.S. Census Bureau. TIGER/Line Shapefiles, a. URL https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html.

U.S. Census Bureau. County Business Patterns (CBP) Data, b. URL https://www.census.gov/programs-surveys/cbp/data.html.

U.S. Geological Survey. The National Map - Data Delivery. URL https://www.usgs.gov/the-national-map-data-delivery.

XGBoost Documentation. XGBoost Documentation: Notes on Parameter Tuning, 2025. URL https://xgboost.readthedocs.io/en/stable/tutorials/param_tuning.html.

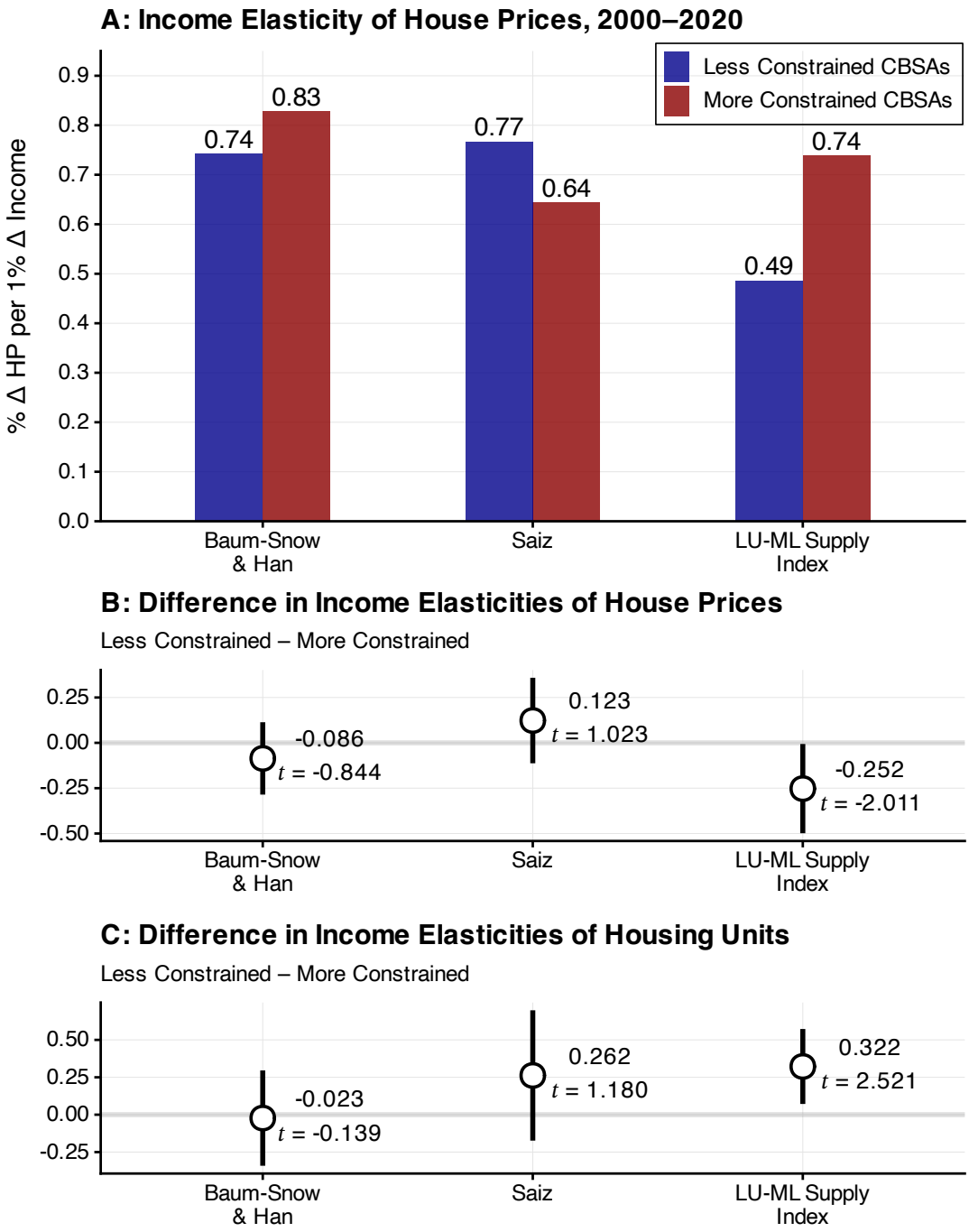
Zillow. Zillow Housing Data. URL https://www.zillow.com/research/data/.

Figure 1: Predictive Power for House Prices: LU-ML Price Pressure vs Saiz and BSH



Notes: This figure assesses the predictive power of supply constraint measures for house prices using weak instrument tests (first-stage F-statistics). We compare the Saiz (2010) and Baum-Snow and Han (BSH, 2024) elasticity proxies against the LU-ML Price Pressure Index. The LU-ML Price Pressure Index is constructed using house prices as the target variable (in levels or differences, matching the respective specification). The gray-dashed line marks the traditional weak instrument threshold of 10 (Stock and Yogo, 2005); the black-dashed line marks a modern threshold of 50 suggested by Keane and Neal (2024). Panel A replicates leading studies. Chaney et al. (2012) use a panel (1993–2007) of MSA-level log house prices, instrumenting with the interaction of the Saiz elasticity and national mortgage rates. Mian and Sufi (2014) use a long-difference specification (2006–2009) for counties. Stroebel and Vavra (2019) use long differences (2001–2007, 2007–2011) for MSAs. Guren et al. (2021) use a quarterly panel (1978–2017) of year-over-year (YoY) log price changes for CBSAs, interacting the Saiz proxy with the change in national house prices. Panels B and C report results for long-difference and panel models, respectively, using Freddie Mac house price indices. In Panel B, standard errors are clustered by commuting zone. Panel C reports results for the panel specification defined in Equation (1), where the dependent variable is the YoY log first-difference in house prices. We interact the BSH and Saiz elasticity proxies with the YoY log first-difference in national house prices, following Guren et al. (2021), to generate time-varying proxies. Standard errors are clustered by commuting zone and time. See Appendix Tables A1 through A7 for full regression output.

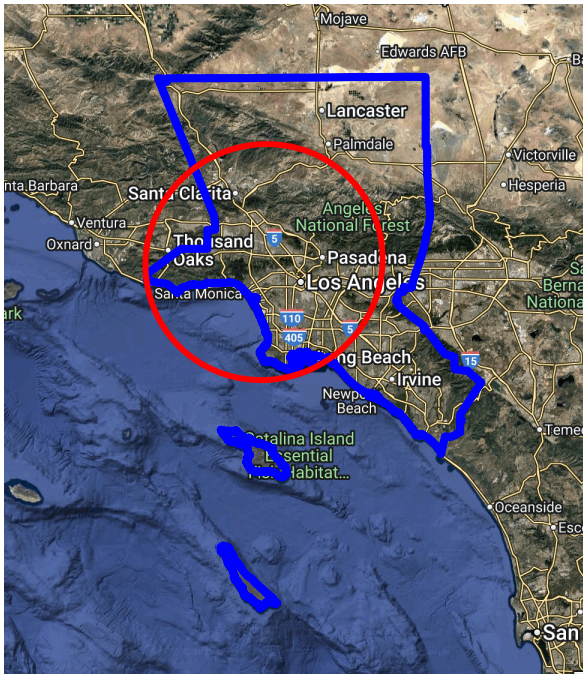
Figure 2: Impact of Income on House Prices by Housing Supply Measure



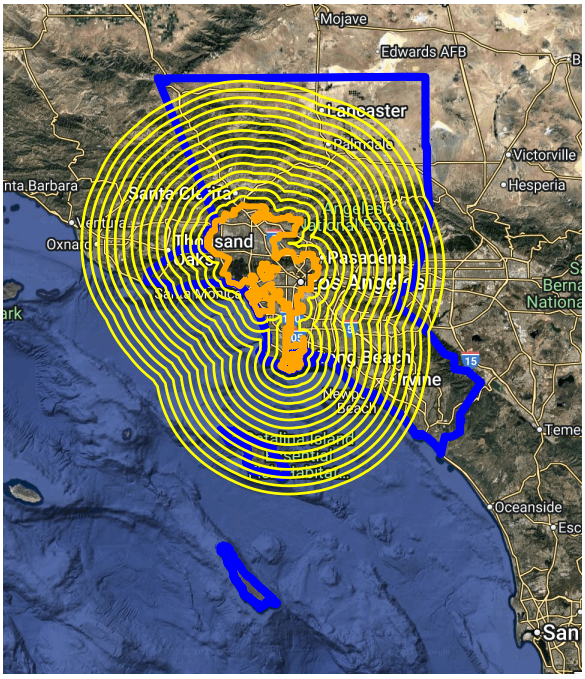
Notes: This figure tests whether looser supply constraints mitigate house price appreciation in response to income growth. We compare results using the Baum-Snow and Han (2024) and Saiz (2010) proxies against the LU-ML Supply Index. The LU-ML Supply Index is constructed using the log change in CBSA-level housing units from 2000–2020 as the target variable. Estimates are derived from the following specification: $\Delta \log(Y_i) = \alpha + \beta_1 \Delta \log(Income_i) + \beta_2 LessConstrained_i + \beta_3 [\Delta \log(Income_i) \times LessConstrained_i] + \epsilon_i$. $LessConstrained_i$ is an indicator equal to 1 for CBSAs with looser-than-average supply constraints according to the respective proxy. Panel A plots the implied income elasticity of house prices (β_1 versus $\beta_1 + \beta_3$) for *More* versus *Less Constrained* CBSAs. Panel B plots the difference in elasticities (β_3), testing if looser constraints dampen price growth. Panel C plots the difference in the income elasticity of housing units, testing if looser constraints facilitate greater quantity expansion. The dependent variable (Y_i) is the Freddie Mac house price index in Panels A and B, and the number of housing units (aggregated from Census block groups via NHGIS) in Panel C. *Income* is defined as total BEA income for Panels A and B and per capita BEA income for Panel C. Changes are calculated from 2000 to 2020, with all data aggregated to 2020 CBSA delineations. t-statistics in Panels B and C are based on robust standard errors clustered at the 2020 commuting-zone level. See Appendix Table A10 for full regression output.

Figure 3: Land Unavailability Circles and Buffers for Los Angeles

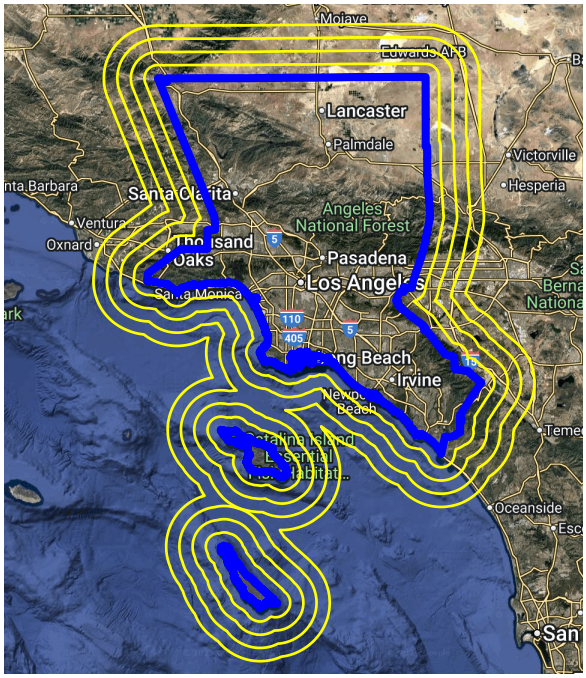
A: LA-Long Beach-Anaheim CBSA with Saiz 50km Circle



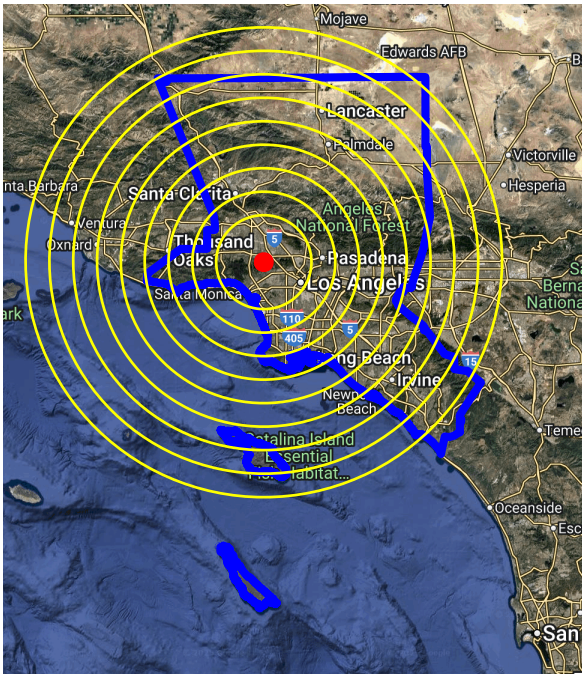
B: LU-ML First Principal City Buffers



C: LU-ML CBSA Polygon Buffers

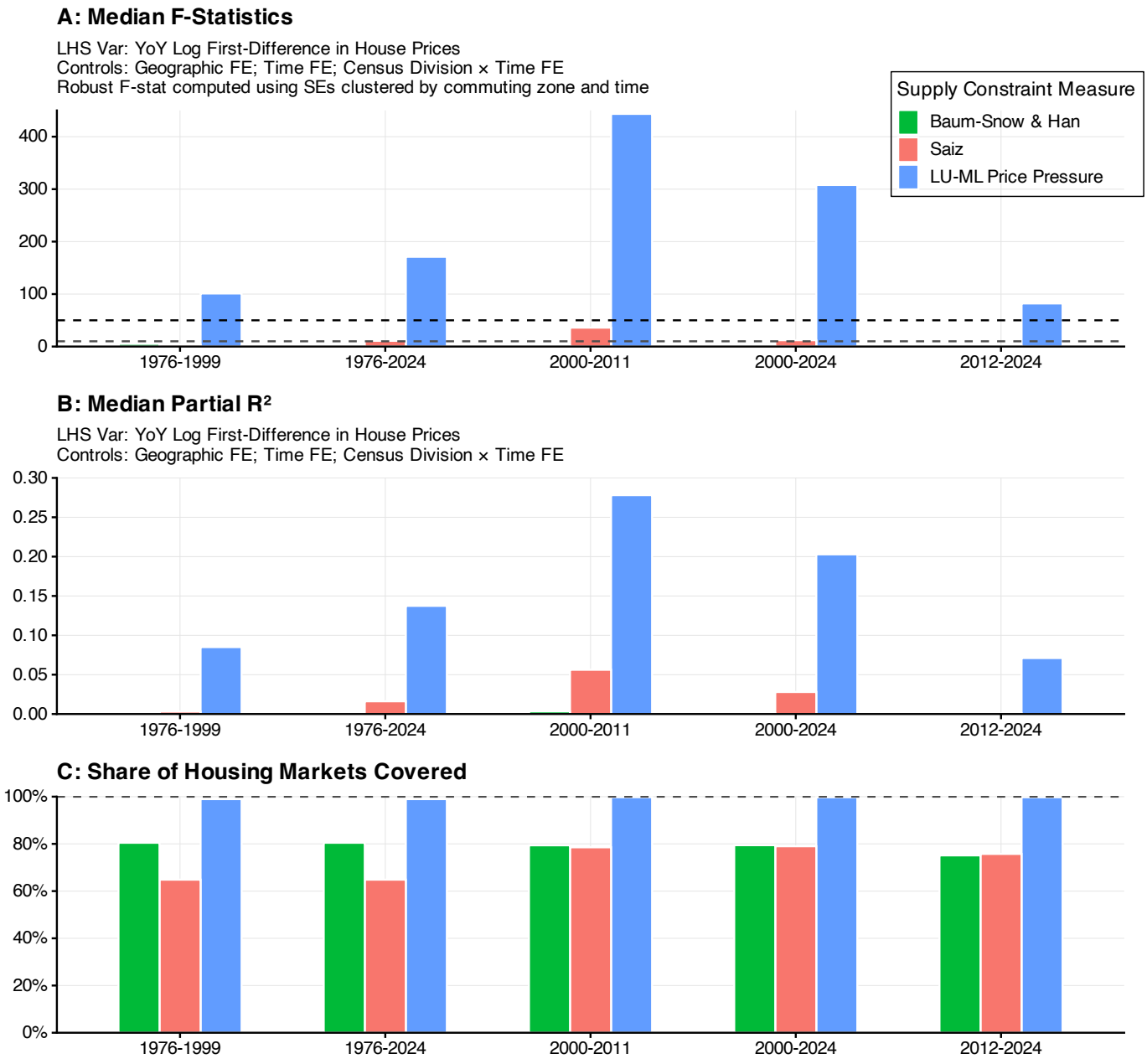


D: LU-ML First Principal City Centroid Circles



Notes: This figure illustrates the spatial definitions used to calculate Land Unavailability. In each panel, the blue boundary represents the Los Angeles-Long Beach-Anaheim CBSA (2023 delineations). Panel A displays the standard definition from Saiz (2010) and Baum-Snow and Han (2024): a 50km radius circle (red) centered on the centroid of the first principal city (Los Angeles). Panels B–D illustrate the varying spatial features used in the LU-ML indices. Panel B displays the first principal city polygon (orange) and its surrounding buffers (yellow) ranging from 10 to 150 percent (in 10pp increments). Panel C displays the CBSA polygon (blue) and its surrounding buffers (yellow) ranging from 5 to 20 percent (in 5pp increments). Panel D displays concentric circles with radii ranging from 20km to 100km (in 10km increments) centered on the first principal city centroid (red dot).

Figure 4: Predictive Power for House Prices in Panels: LU-ML vs Saiz and BSH



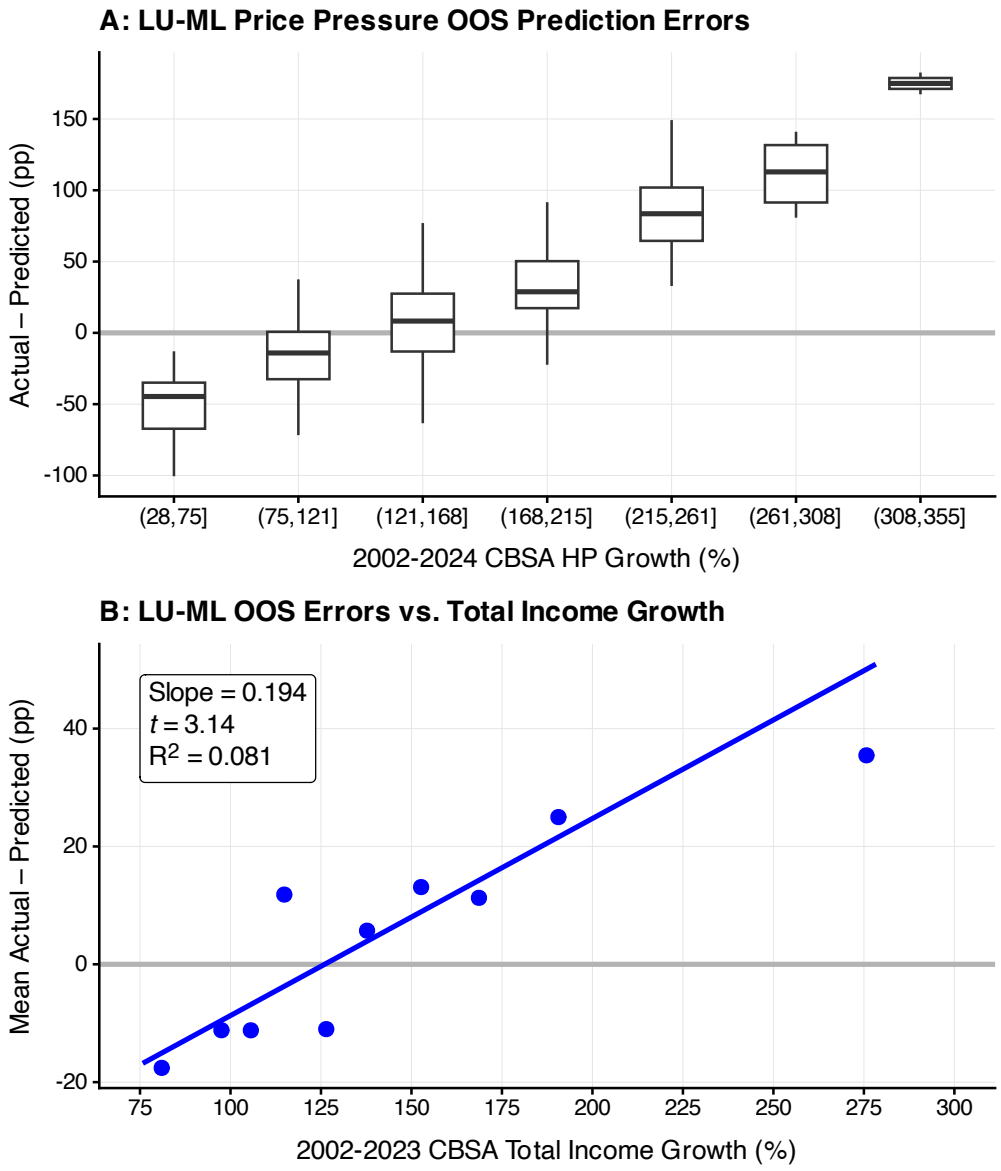
Notes: This figure summarizes the predictive power of supply constraint measures across a comprehensive set of housing panel datasets. We compare the Saiz (2010) and Baum-Snow and Han (2024) elasticity proxies against the LU-ML Price Pressure Index. The summary includes panel datasets from Freddie Mac, FHFA, and Zillow, aggregated to various geographic levels including CBSAs, counties, three-digit zip codes, zip codes, and census tracts. See Tables A8 and A9 for the full list of specifications. Estimates are derived from the following specification: $\Delta \log HP_{i,d,t} = \psi_i + \gamma_{d,t} + \beta \cdot SC_{i,d,t} + \varepsilon_{i,d,t}$. $\Delta \log HP_{i,d,t}$ is the year-over-year (YoY) log first-difference in house prices for geographic unit i (e.g., a zip code or CBSA) in Census Division d at month t. ψ_i denotes geographic unit fixed effects, and $\gamma_{d,t}$ denotes Census division-by-time fixed effects. $SC_{i,d,t}$ represents the supply constraint measure. To generate time-varying instruments for the static Saiz and BSH proxies, we interact them with the YoY log change in national house prices, following Guren et al. (2021). The LU-ML Price Pressure Index is constructed for each specific dataset using the corresponding YoY price change as the target variable. Panel A reports weak instrument tests—the median first-stage F-statistic across all specifications within each estimation period. The gray-dashed line marks the weak instrument threshold of 10 (Stock and Yogo, 2005); the black-dashed line marks the threshold of 50 (Keane and Neal, 2024). Panel B reports the median first-stage partial R^2. Panel C reports the coverage rate: the share of available cross-sectional geographic units for which a supply constraint measure exists, averaged across specifications in that period.

Table 1: Robustness of LU-ML Price Pressure Estimates to Time-Varying Natural Amenity Demand (2002–2024)

	Dependent variable:				
	YoY Log Diff in House Prices				
	(1)	(2)	(3)	(4)	(5)
Panel A: Freddie Mac House Prices					
LU-ML Price	0.699***	0.659***	0.525***	0.512***	0.476***
Pressure Index	(0.061)	(0.059)	(0.050)	(0.050)	(0.047)
Panel B: Zillow House Prices					
LU-ML Price	0.676***	0.634***	0.520***	0.517***	0.503***
Pressure Index	(0.061)	(0.057)	(0.054)	(0.055)	(0.049)
Panel C: FHFA House Prices					
LU-ML Price	0.665***	0.630***	0.574***	0.566***	0.530***
Pressure Index	(0.065)	(0.067)	(0.057)	(0.057)	(0.054)
CBSA Fixed Effects	✓	✓	✓	✓	✓
Census Division × Date FE	✓	✓	✓	✓	✓
Coastal Dummy × Date FE		✓	✓	✓	✓
Mean Jan Temp Quintiles × Date FE			✓	✓	✓
Mean Jul Temp Quintiles × Date FE			✓	✓	✓
Major Park Dummy × Date FE				✓	✓
GMaps Hiking Demand Quintiles × Date FE					✓
GMaps Water Demand Quintiles × Date FE					✓
GMaps Natural Recreation Quintiles × Date FE					✓

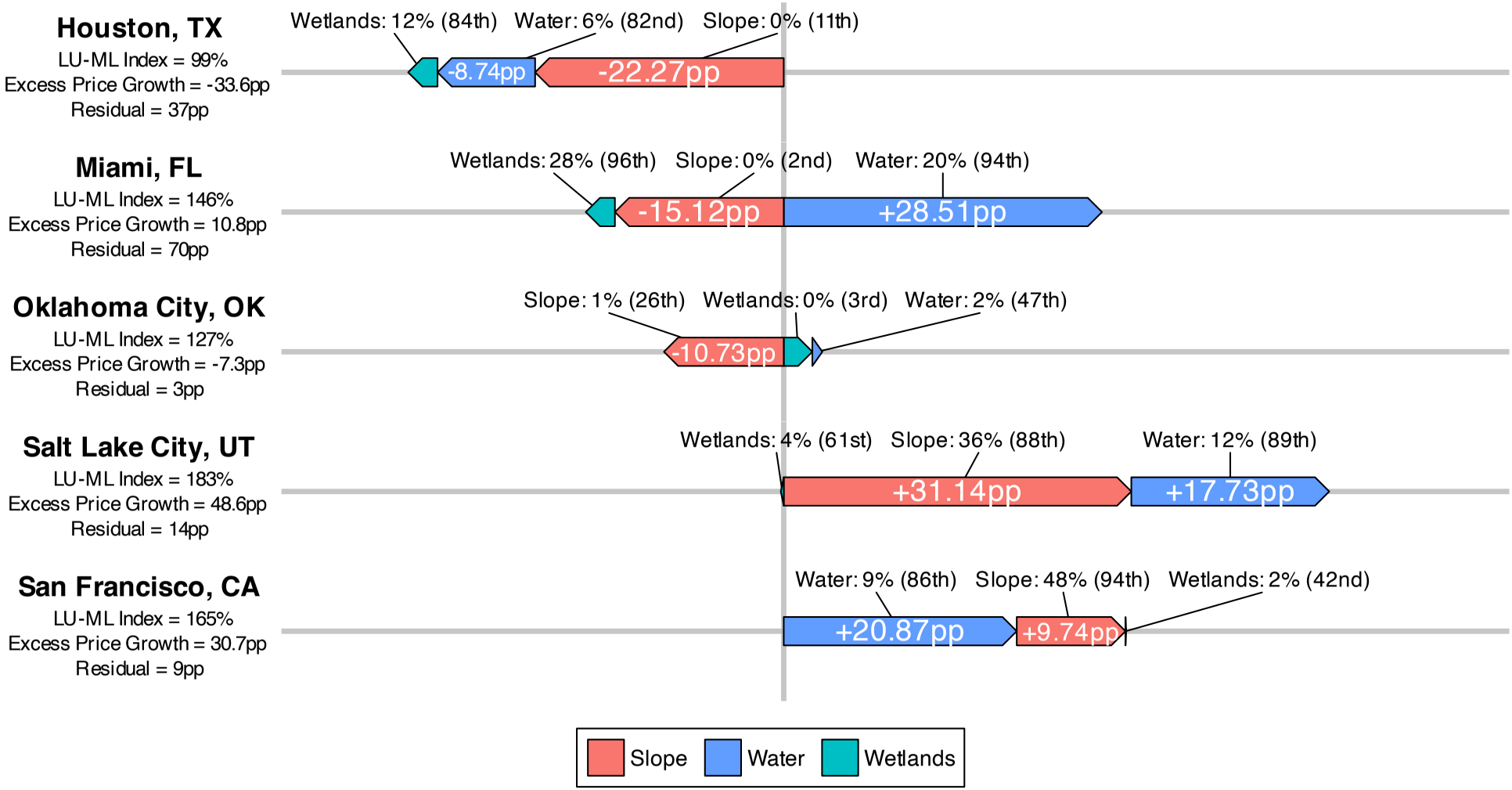
Notes: This table reports panel regressions of year-over-year (YoY) house price growth on the LU-ML Price Pressure Index (2002–2024). The LU-ML Price Pressure Index is constructed separately for each panel (Freddie Mac, Zillow, FHFA) using the respective YoY log price difference as the target variable. Column (1) estimates the baseline specification: $\Delta \log HP_{i,d,t} = \psi_i + \gamma_{d,t} + \beta \cdot LU\text{-}ML_{i,d,t} + \varepsilon_{i,d,t}$. $\Delta \log HP_{i,d,t}$ is the YoY log first-difference in house prices for CBSA i in Census Division d at month t. ψ_i denotes CBSA fixed effects, and $\gamma_{d,t}$ denotes Census division-by-date fixed effects. Columns (2)–(5) progressively add nonparametric controls for time-varying amenity demand by interacting static geographic features with date fixed effects. Column (2) adds a coastal indicator (U.S. Census shapefile) × date fixed effects. Column (3) adds average January and July temperature quintiles (aggregated from county data via NHGIS (Schroeder et al., 2025)) × date fixed effects. Column (4) adds a major park indicator (U.S. National Park Service shapefile) × date fixed effects. Column (5) controls for the Lutz (2025b) Google Maps (GMaps) indices for hiking, water, and natural recreation demand; we interact the quintiles of each index with date fixed effects. Panels A–C use Freddie Mac, Zillow, and FHFA house price indices, respectively. Robust standard errors clustered by commuting zone and time. One, two, or three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Figure 5: LU-ML Prediction Errors and Housing Demand Growth



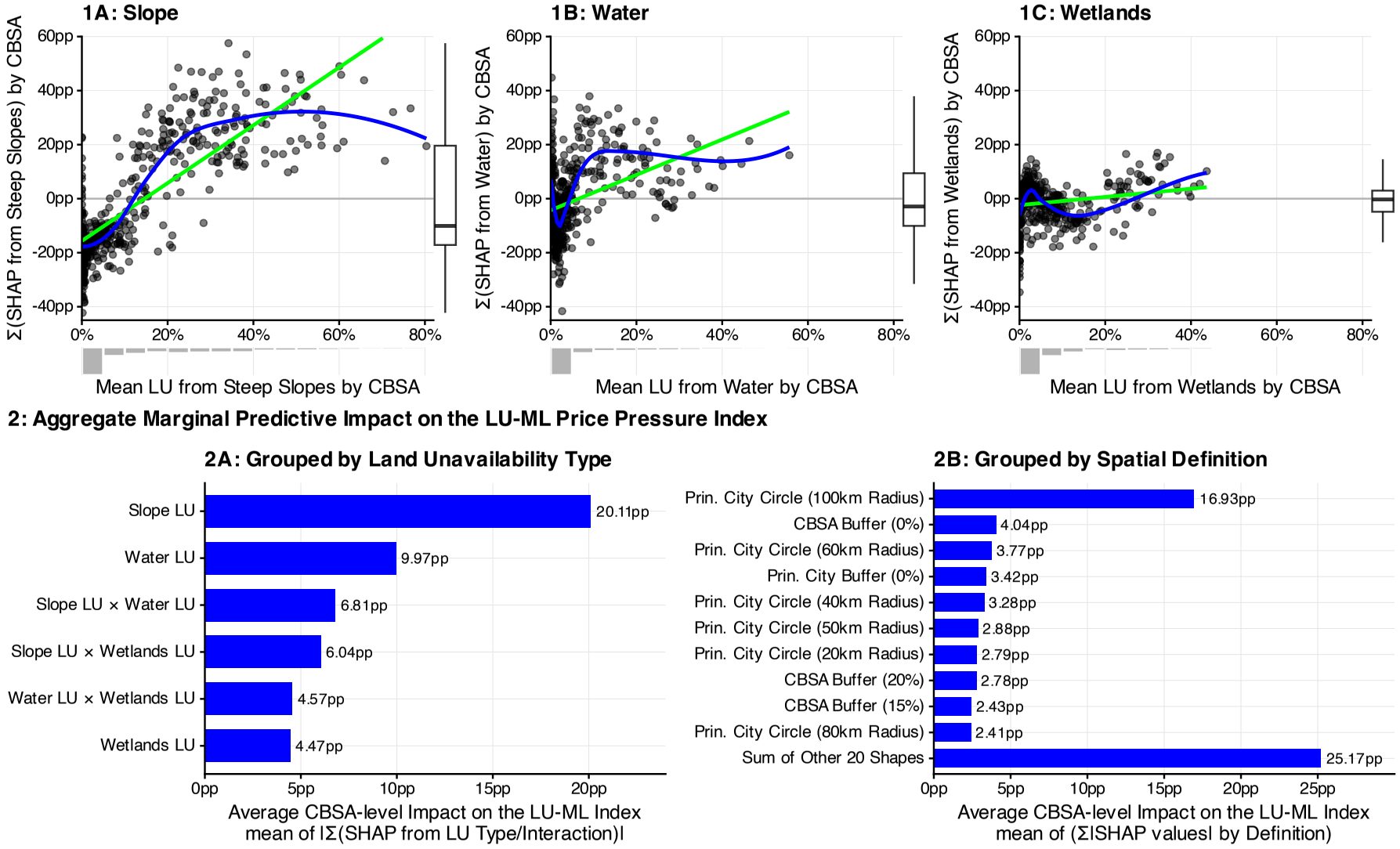
Notes: This figure analyzes the out-of-sample (OOS) prediction errors of the LU-ML Price Pressure Index. Panel A displays boxplots of the OOS prediction errors (vertical axis) sorted by cumulative Zillow CBSA house price growth (horizontal axis) from 2002–2024. We construct the LU-ML Price Pressure Index using 2002–2024 Zillow cumulative house price growth as the target variable. Panel B presents a binscatter of the OOS prediction errors against 2002–2023 BEA total income growth. We select the number of bins following Cattaneo et al. (2024). The regression line and statistics are estimated on the underlying data; standard errors are clustered by commuting zone (2020 definitions).

Figure 6: Decomposing the LU-ML Price Pressure Index—Feature Contributions in Representative Cities



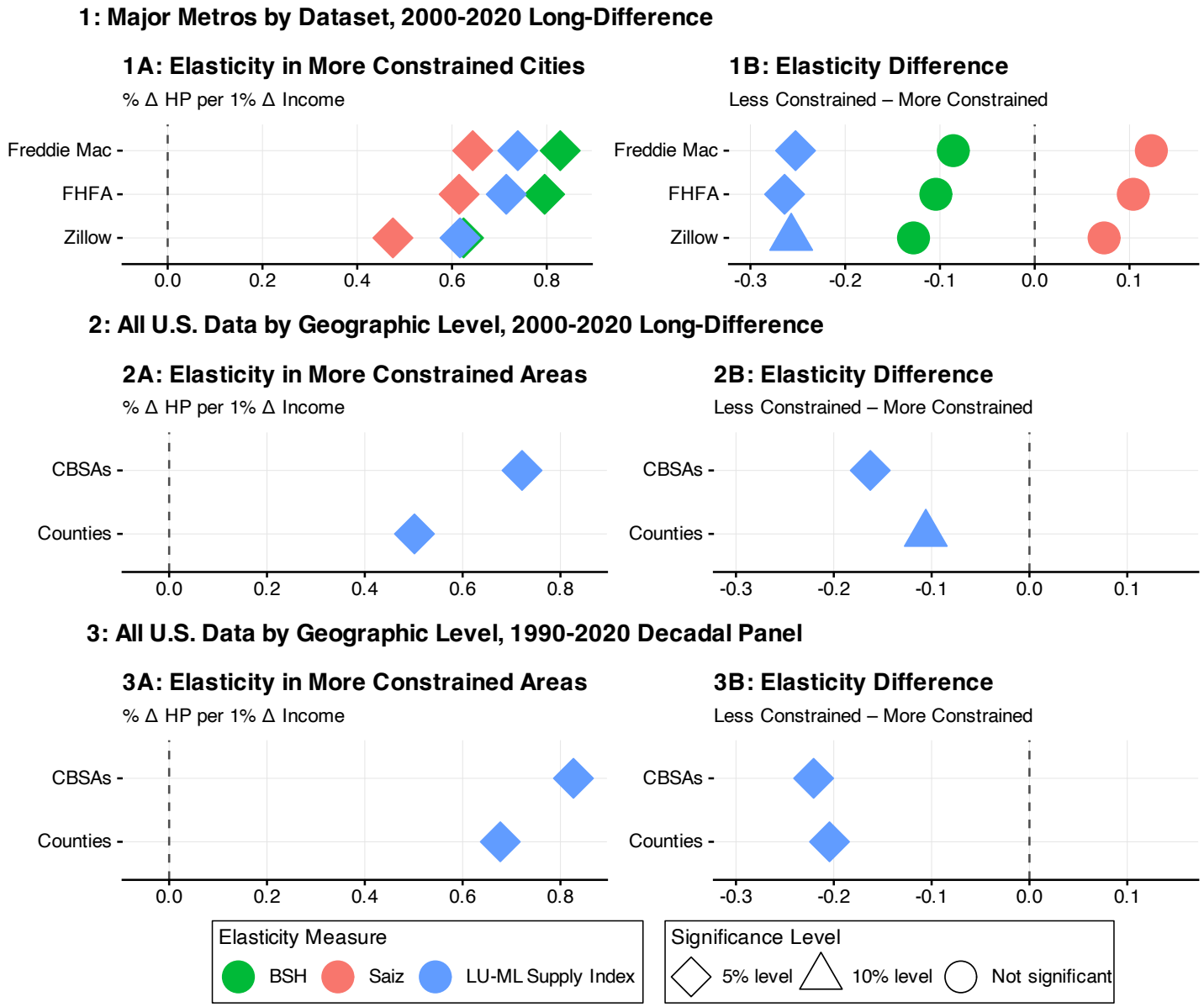
Notes: This figure decomposes the LU-ML Price Pressure Index for representative CBSAs using SHAP values (Lundberg and Lee, 2017; Lundberg et al., 2020). The LU-ML Price Pressure Index is trained using 2002–2024 Zillow cumulative house price growth as the target variable and employs 90 total features (slope, water, and wetlands LU measured using 30 different spatial definitions). The SHAP value, $\varphi_i^{(j)}$, measures the marginal contribution of feature j to the LU-ML Index for CBSA i relative to a coalition of other features ($\neq j$), according to a fair distribution of payouts from cooperative game theory (Shapley, 1953). This allows for an additive decomposition of the index: LU-ML Index$_i = \mu + \sum_{j=1}^{M} \varphi_i^{(j)}$, where μ represents the unconditional mean prediction across all CBSAs. Text below each city label reports the total predicted growth (LU-ML Index), the predicted growth in excess of the national mean ($\sum_j \varphi^{(j)}$), and the prediction error (Residual). The arrows visualize the SHAP contributions: we aggregate the SHAP values from the 90 underlying features into three categories—Slope (red), Water (blue), and Wetlands (teal). The length of each arrow represents the feature set's contribution to the deviation from the mean prediction (μ), measured in percentage points (pp). Labels on the arrows report the underlying land unavailability percentages for that category, with national percentiles in parentheses.

Figure 7: Global Feature Contributions and Nonlinearities in the LU-ML Price Pressure Index (2002–2024)



Notes: This figure summarizes the marginal predictive impact of LU features (variables) on the LU-ML Price Pressure Index. The LU-ML Price Pressure Index is trained using 2002–2024 cumulative house price growth as the target variable and employs 90 total features (slope, water, and wetlands LU measured using 30 different spatial definitions). We calculate marginal predictive impacts using SHAP values (Lundberg and Lee, 2017; Lundberg et al., 2020). The SHAP value, $\varphi_i^{(j)}$, measures the marginal contribution of LU feature j to the LU-ML Index for CBSA i relative to a coalition of other features ($\neq j$), according to a fair distribution of payouts from cooperative game theory (Shapley, 1953). Panels 1A–1C illustrate aggregate nonlinearities. For each CBSA, we sum the SHAP values associated with a specific LU type (Slope, Water, or Wetlands) and plot this net contribution (vertical axis) against the mean LU share for that type (horizontal axis). The fitted lines show the average marginal impact across the distribution of land unavailability. Panel 2 compares feature importance using two different aggregation methods. Panel 2A reports the average magnitude of the *net* impact for each LU type or interaction. Specifically, we sum the SHAP values for all features within a category (e.g., Slope) to find the net impact for each CBSA, then average the absolute values of these sums across all CBSAs. Panel 2B reports the *gross* predictive contribution of each spatial definition (polygon shape). Specifically, we sum the absolute SHAP values for all features derived from a specific polygon (e.g., Principal City 100km Radius) within each CBSA, then average these sums across all CBSAs.

Figure 8: Supply Constraints and the Income Elasticity of House Prices



Notes: This figure plots the estimated impact of income growth on house price growth for *Less Constrained* versus *More Constrained* areas by supply constraint proxy. *Less Constrained* areas are defined as those with looser-than-average supply constraints. Supply proxies are from Baum-Snow and Han (2024, BSH), Saiz (2010), and the LU-ML Supply Index. We construct the LU-ML Supply Index using the log change in housing units (2000–2020) as the target variable. Panels 1 and 2 show long-difference estimates from 2000–2020 using the following regression:

$$\Delta \log(HP_i) = \alpha + \beta_1 \Delta \log(Income_i) + \beta_2 LessConstrained_i + \beta_3 [\Delta \log(Income_i) \times LessConstrained_i] + \epsilon_i$$

We estimate separate regressions using each supply constraint proxy (BSH, Saiz, LU-ML). Panel 1 limits the data to major metropolitan areas (Freddie Mac sample). Panel 2 expands the sample to all U.S. CBSAs and counties using FHFA data; BSH and Saiz are omitted due to limited geographic coverage. Panel 3 reports estimates from a decadal panel (1990–2020) using all U.S. data (FHFA):

$$\Delta \log(HP_{it}) = \alpha_i + \eta_t + \beta_1 \Delta \log(Income_{it}) + \beta_3 [\Delta \log(Income_{it}) \times LessConstrained_i] + \epsilon_{it}$$

where α_i and η_t represent geographic and decadal-time fixed effects, respectively. In all panels, the left column (A) plots β_1 (elasticity in More Constrained areas), and the right column (B) plots β_3 (the difference in elasticity). Robust standard errors are clustered by commuting zone (2020 definitions). Regression output is available in Table A11.

Table 2: Income Elasticity of House Prices – Interactions with Continuous LU-ML Supply Index (1990–2020 CBSA Decadal Panel)

	Dependent variable:			
	$\Delta \log$ HP			
	(1)	(2)	(3)	(4)
Total Income Growth	0.624^{***}	0.555^{***}	0.629^{***}	0.737^{***}
	(0.031)	(0.032)	(0.039)	(0.048)
LU-ML Supply Index (Standardized)	0.046^{***}	0.043^{***}	0.055^{***}	
	(0.010)	(0.010)	(0.012)	
Total Income Growth ×	-0.090^{***}	-0.097^{***}	-0.118^{***}	-0.150^{***}
LU-ML Supply Index (Standardized)	(0.022)	(0.022)	(0.025)	(0.032)
Decadal-Time FE	✓	✓	✓	✓
State FE		✓		
Commuting Zone FE			✓	
CBSA FE				✓

Notes: This table reports the estimated elasticity of house prices with respect to income, interacting income growth with the continuous LU-ML Supply Index. We construct the LU-ML Supply Index using the log change in housing units (2000–2020) as the target variable. The LU-ML Supply Index is standardized to have mean zero and unit variance. Estimates are based on the following specification:

$$\Delta \log(HP_{it}) = \alpha_i + \eta_t + \beta_1 \Delta \log(Income_{it}) + \beta_2 LU\text{-}ML\ Index_i + \beta_3 [\Delta \log(Income_{it}) \times LU\text{-}ML\ Index_i] + \epsilon_{it}$$

where η_t represents decadal-time fixed effects and α_i signifies geographic fixed effects (State, Commuting Zone, or CBSA depending on the column). The coefficient β_1 represents the estimated elasticity for a location with average supply constraints. The interaction coefficient β_3 estimates the change in elasticity associated with a one standard deviation increase in the supply index (indicating less constrained supply). The main effect of the time-invariant supply index is absorbed by the CBSA fixed effects in Column (4). Robust standard errors are clustered by commuting zone (2020 definitions). One, two, and three asterisks represent statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 3: Income Elasticity of House Prices – Interactions with the Time-Varying LU-ML Supply Index (1990–2020 Decadal Panel)

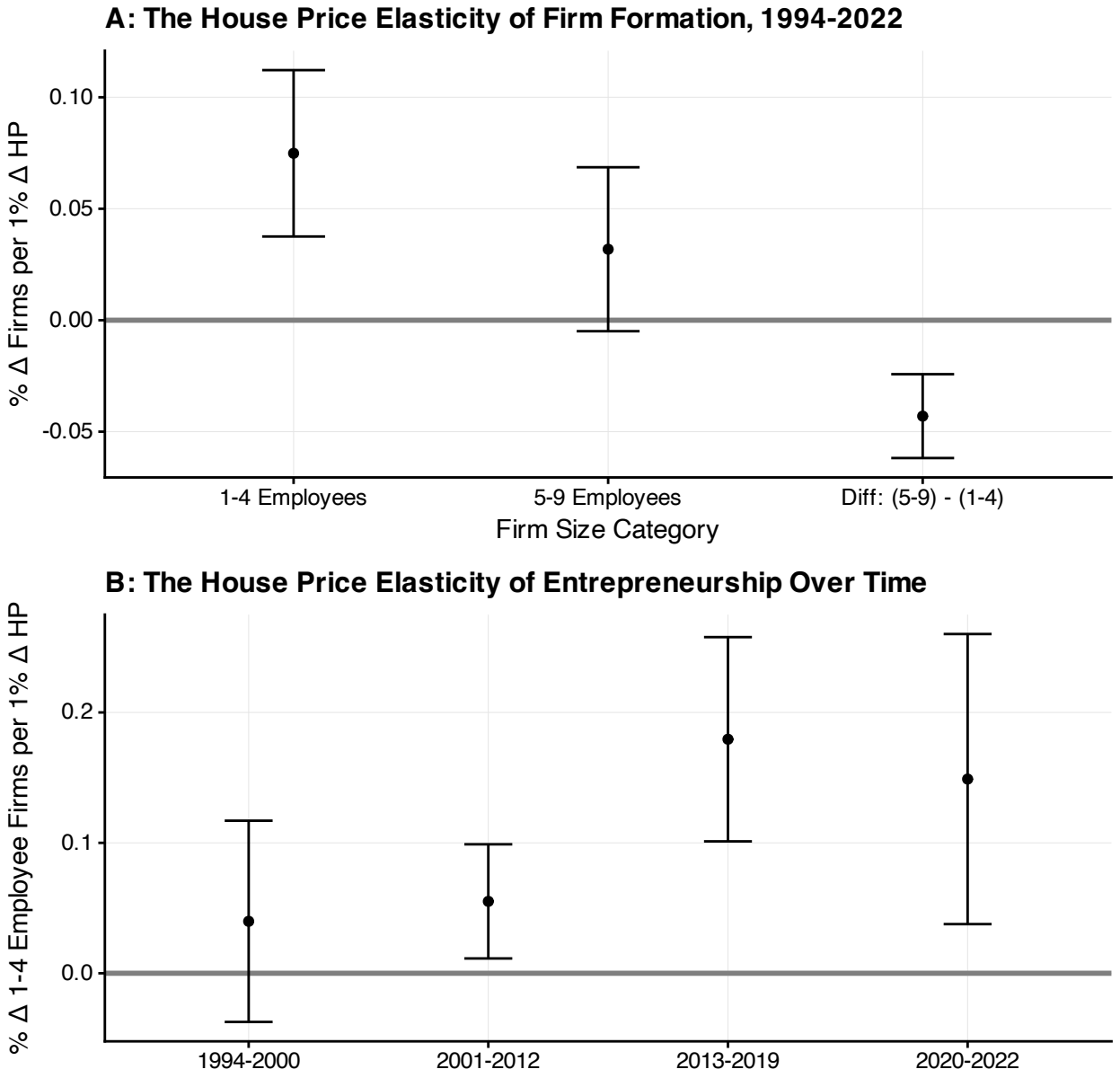
	Dependent Var: $\Delta \log HP$	
	CBSAs	Counties
	(1)	(2)
Total Income Growth	0.734^{***}	0.608^{***}
	(0.045)	(0.051)
LU-ML Supply Index (Std. by Decade)	0.071^{***}	0.044^{***}
	(0.012)	(0.015)
Total Income Growth ×	-0.140^{***}	-0.069^{**}
LU-ML Supply Index (Std. by Decade)	(0.027)	(0.029)
Decadal-Time FE	✓	✓
CBSA/County FE	✓	✓

Notes: This table reports the estimated elasticity of house prices with respect to income, interacting income growth with the time-varying panel LU-ML Supply Index. We construct the panel LU-ML Supply Index separately for each decade (1990s, 2000s, 2010s) using the log change in housing units as the target variable. The Index is standardized within each decade to have zero mean and unit variance. Estimates are based on the following specification:

$$\Delta \log(HP_{it}) = \alpha_i + \eta_t + \beta_1 \Delta \log(Income_{it}) + \beta_2 LU\text{-}ML\ Index_{it} + \beta_3[\Delta \log(Income_{it}) \times LU\text{-}ML\ Index_{it}] + \epsilon_{it}$$

Each column reports results from a separate regression using county or CBSA decadal panel data (1990–2020). All regressions include geographic (CBSA or county) and decadal-time fixed effects. Robust standard errors are clustered by commuting zone (2020 definitions). One, two, and three asterisks represent statistical significance at the 10, 5, and 1 percent levels, respectively.

Figure 9: The Impact of House Price Growth on Entrepreneurship (2SLS Estimates, 1994–2022)



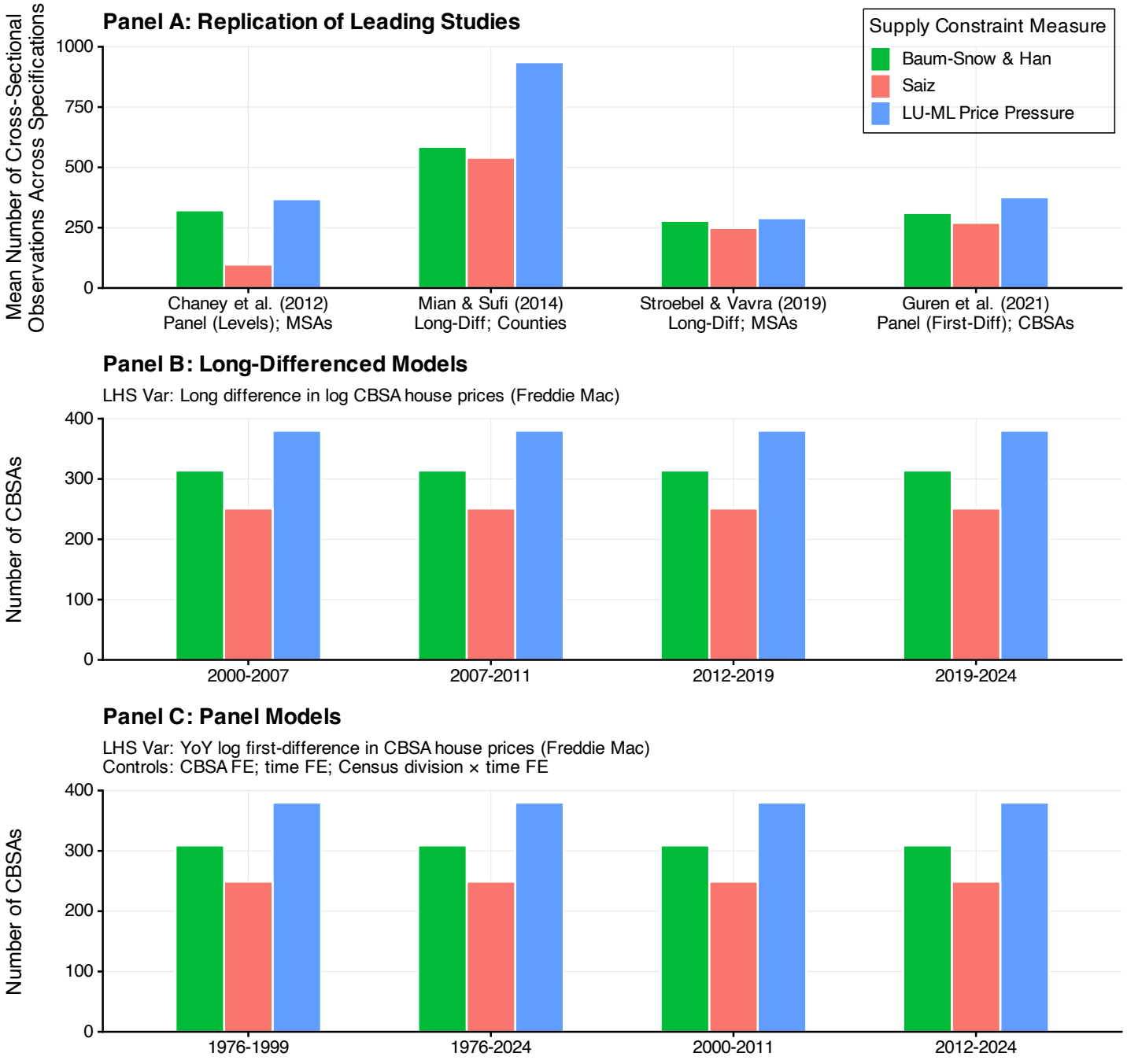
Notes: This figure plots the impact of house price growth on entrepreneurship, proxied by the growth in firms with 1–4 employees. Estimates are derived from the following two-stage least squares (2SLS) specification:

$$\Delta \log(\text{Firms}_{i,s,t}^{1-4}) = \alpha_i + \delta_{s,t} + \beta \cdot \Delta \log(\text{HP}_{i,s,t}) + \epsilon_{i,s,t}$$

where $\Delta \log(\text{Firms}_{i,s,t}^{1-4})$ is the year-over-year (YoY) log first-difference in the number of firms with 1–4 employees in zip code i and state s at year t. α_i represents zip code fixed effects, $\delta_{s,t}$ signifies state-time fixed effects, and $\Delta \log(\text{HP}_{i,s,t})$ is the YoY log difference in zip code FHFA house prices. We instrument house price growth with the zip-code-level LU-ML Price Pressure Index. We construct the LU-ML Price Pressure Index using $\Delta \log(\text{HP}_{i,s,t})$ as the target variable. Panel A plots β for 1–4 employee firms, 5–9 employee firms, and the difference between 1–4 and 5–9 employee firms. Panel B plots the estimates from an augmented specification where we interact house prices (and the LU-ML instrument) with period indicators (1994–2000, 2001–2012, 2013–2019, 2020–2022). Robust standard errors are clustered by 2020 commuting zone and time. Error bars correspond to ±2 standard errors.

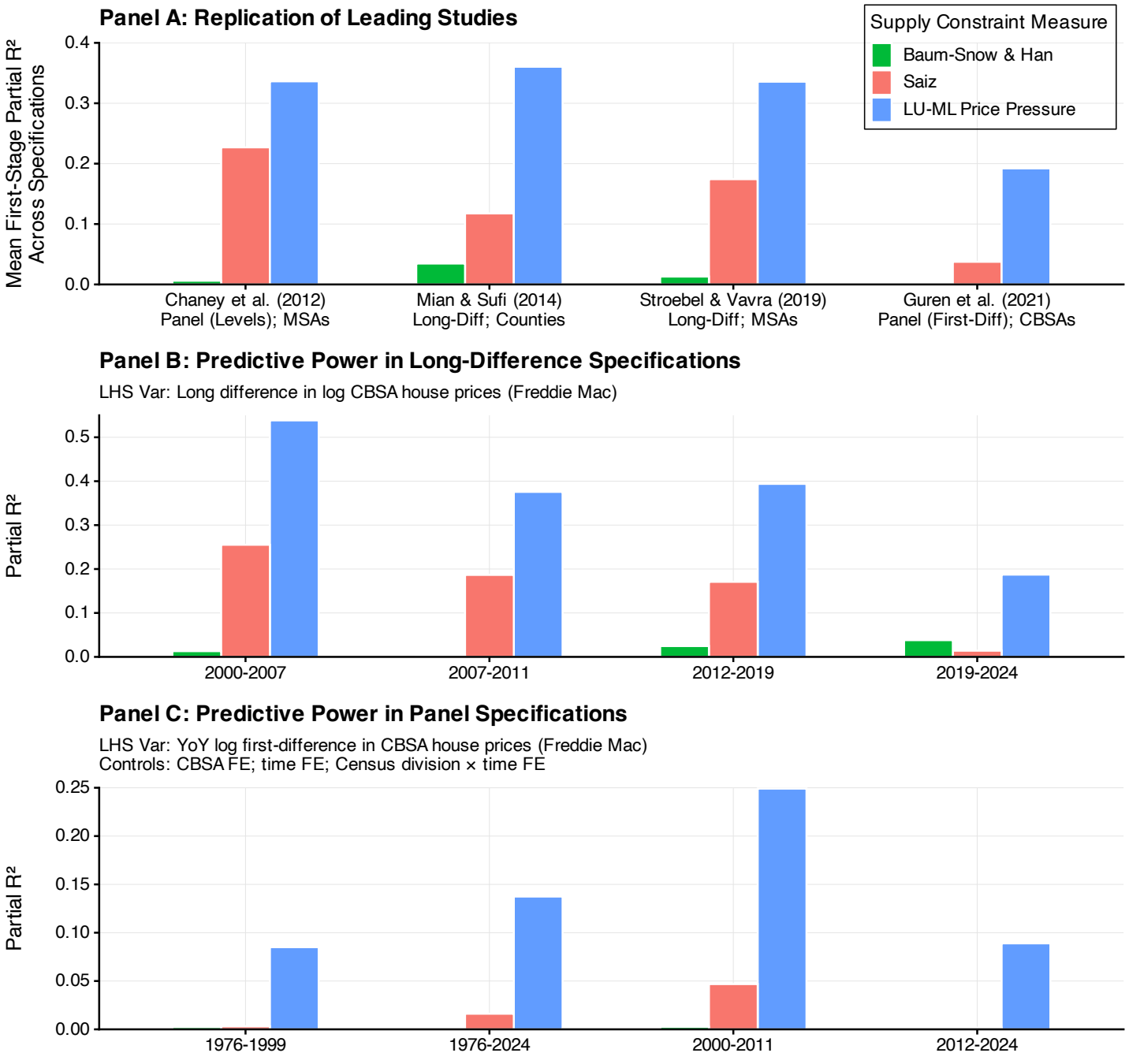
A Appendix: The Predictive Power of Supply Constraints

Figure A1: Cross-Sectional Observation Counts: LU-ML vs Saiz and BSH



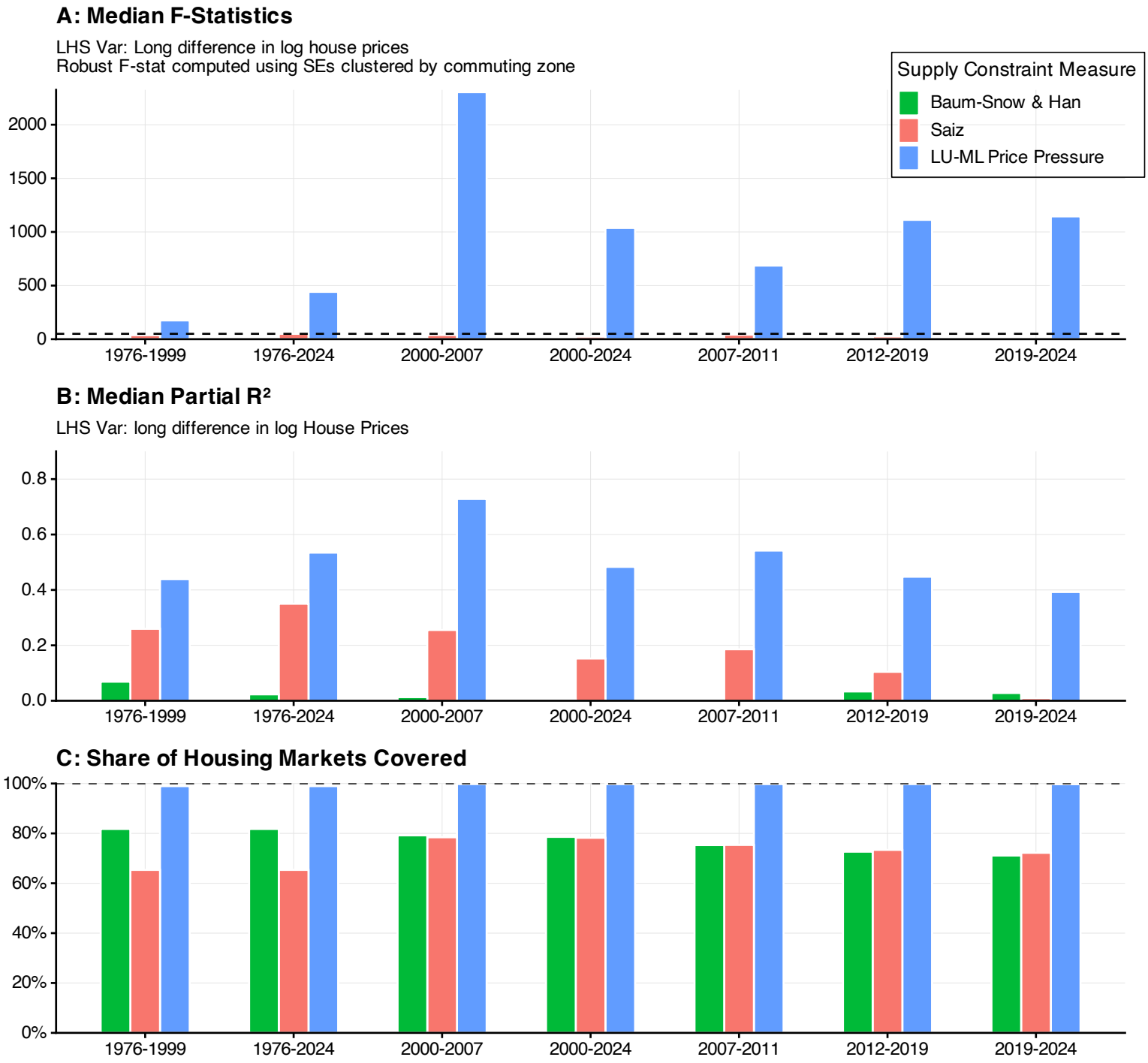
Notes: See the notes to Figure 1.

Figure A2: Partial R^2 Statisitics: LU-ML vs Saiz and BSH



Notes: See the notes to Figure 1.

Figure A3: Predictive Power in Long-Difference Specifications: LU-ML vs Saiz and BSH



Notes: This figure summarizes the predictive power of supply constraint measures for house prices in long difference specifications across a comprehensive set of housing panel datasets. We compare the Saiz (2010) and Baum-Snow and Han (2024) elasticity proxies against the LU-ML Price Pressure Index. The summary includes panel datasets from Freddie Mac, FHFA, and Zillow, aggregated to various geographic levels including CBSAs, counties, three-digit zip codes, zip codes, and census tracts. Panel A reports weak instrument tests—the median first-stage F-statistic across all specifications within each estimation period. The black-dashed horizontal line represents a first-stage F-statistic of 50, as recommended by Keane and Neal (2024). Panel B reports the median first-stage partial R^2. Panel C reports the coverage rate: the share of available cross-sectional geographic units for which a supply constraint measure exists, averaged across specifications in that period.

Table A1: First Stage House Price Regression as in Chaney et al. (2012)

	Dependent variable:		
	MSA HP Residential Prices		
	(1)	(2)	(3)
Saiz Elasticity × Natl Mortgage Rate	0.029***		
	(0.005)		
Baum-Snow & Han × Natl Mortgage Rate		0.023	
		(0.017)	
LU-ML			0.780***
			(0.038)
First Stage F-Stat	40.23	1.79	417.82
First Stage Partial R^2	0.23	0.01	0.34
Number of MSAs	97	322	368
Observations	1,358	4,499	5,132
MSA Fixed Effects	✓	✓	✓
Time Fixed Effects	✓	✓	✓

Notes: Column (1) replicates the first stage regression in Table 3, column (1) of Chaney et al. (2012) that uses Saiz Elasticity as an instrument for house prices using their equation (2): $P_t^l = \alpha^l + \delta_t + \gamma \cdot Elasticity^l \times IR_t + u_t^l$. P_t^l is the normalized residential house price index (in levels) for MSA l in year t, α^l and δ_t are MSA and time fixed effects, $Elasticity^l \times IR_t$ is the Saiz Elasticity proxy for each MSA multiplied by the national real mortgage rate, and u_t^l is the error term. Column (2) uses our LU-ML instrument for 2007 MSAs instead of $Elasticity^l \times IR_t$. The sample period ranges from 1993 to 2007. Robust standard errors clustered at the MSA level are in parentheses. One, two, or three asterisks represent statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A2: Replication of Mian and Sufi (2014) – Non-Tradable Employment and the Housing Net Worth

	Non-Tradable Emp. Growth, 2007-09							
	Rest. & Retail	Geog. Concen.	Rest. & Retail	Geog. Concen.	Rest. & Retail	Geog. Concen.	Rest. & Retail	Geog. Concen.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Housing Net Worth, 2006-09	0.174***	0.166***	0.374***	0.208**	0.160	−0.073	0.223***	0.226***
	(0.022)	(0.017)	(0.132)	(0.086)	(0.169)	(0.151)	(0.055)	(0.051)
First Stage F-Stat			11.07	11.07	3.49	3.49	241.68	241.68
First Stage Partial R^2			0.12	0.12	0.03	0.03	0.36	0.36
Specification	OLS	OLS	IV	IV	IV	IV	IV	IV
Instrument			Saiz Elasticity	Saiz Elasticity	Baum-Snow & Han	Baum-Snow & Han	LU-ML	LU-ML
Number of Counties	944	944	540	540	585	585	936	936

Notes: Columns 1 to 4 replicate Mian and Sufi (2014) using the equation $\Delta \log E_i^{NT} = \alpha + \eta \cdot \Delta HNW_i + \varepsilon_i$. $\Delta \log E_i^{NT}$ is the log change in non-tradable employment for county i and HNW_i is the change in housing net worth for county i. Mian and Sufi proxy non-tradable employment via the restaurant and retail sector (*Rest. and Retail*) or through industries that have low geographic concentration (*Geog. Concen.*). Columns (5) to (6) use the elasticity proxy from Baum-Snow and Han (2024) as an instrument, while columns (7) and (8) employ the LU-ML IV. Controls include 23 two-digit 2006 employment shares. Robust standard errors are clustered by state. One, two, or three asterisks represent statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A3: First Stage House Price Regression as in Stroebel and Vavra (2019)

	House Price Growth					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 2001–06						
Saiz Elasticity	−0.069***	−0.058***				
	(0.012)	(0.011)				
Baum-Snow & Han			0.242***	0.152*		
			(0.084)	(0.083)		
LU-ML					1.027***	0.881***
					(0.064)	(0.066)
Number of MSAs	250	247	280	276	291	287
First Stage F-Stat	34.81	27.29	8.24	3.39	259.63	177.85
First Stage Partial R^2	0.23	0.20	0.03	0.02	0.44	0.35
Controls		✓		✓		✓
Panel B: 2007–11						
Saiz Elasticity	0.069***	0.038***				
	(0.012)	(0.007)				
Baum-Snow & Han			−0.080	−0.038		
			(0.104)	(0.076)		
LU-ML					1.007***	0.660***
					(0.093)	(0.094)
Number of MSAs	250	248	280	277	291	288
First Stage F-Stat	33.87	30.38	0.59	0.26	116.60	49.67
First Stage Partial R^2	0.19	0.08	0.00	0.00	0.34	0.21
Controls		✓		✓		✓

Notes: Columns (1) and (2) replicate the first stage regression in Table A2, columns (1) and (5) of Stroebel and Vavra (2019) using the equation $\Delta \log(HP_m) = \alpha + \beta Elasticity_m + \gamma X_m + \varepsilon_m$. $\Delta \log(HP_m)$ is the log difference in house prices for MSA m between 2001–06 (panel A) or 2007–11 (panel B). Stroebel and Vavra (2019) use proprietary house prices from CoreLogic. This replication uses publicly available Freddie Mac House Price Indices. $Elasticity_m$ is the Saiz elasticity proxy for MSA m. X_m is a vector of controls for MSA m and includes the change in the share of grocery retail employment, the change in the share of nontradable employment, the change in the share of construction employment, the change in the unemployment rate, and the change in the wage. See Table 1 in Stroebel and Vavra (2019) for more information on these controls. The number of MSAs in columns (1) and (2) is higher than in Table A2 of Stroebel and Vavra (2019) because their replication code includes an undisclosed outcome variable that limits the number of MSAs. Without this outcome variable, we could not determine which MSAs were used in their analysis. Here we report regressions results using all available MSAs in the Stroebel and Vavra (2019) data for each specification. Columns (3) and (4) use the elasticity proxy from Baum-Snow and Han (2024) and columns (5) and (6) use our LU-ML instrument. Robust standard errors are in parentheses. One, two, or three asterisks represent statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A4: Replication of Guren et al. (2021), Table 1

	Dependent variable:		
	YoY Log Diff in Retail Emp Per Capita		
	(1)	(2)	(3)
Panel A: 1978–2017			
YoY Log Diff in HP Growth	0.083***	0.058***	0.084*
	(0.007)	(0.017)	(0.048)
First Stage F-Stat		249.08	19.67
First Stage Partial R^2		0.16	0.03
Observations	59,999	59,999	42,710
Panel B: 1990–2017			
YoY Log Diff in HP Growth	0.081***	0.072***	0.141***
	(0.008)	(0.015)	(0.037)
First Stage F-Stat		440.81	20.41
First Stage Partial R^2		0.27	0.04
Observations	41,985	41,985	29,867
Panel C: 2000–2017			
YoY Log Diff in HP Growth	0.068***	0.055***	0.134***
	(0.008)	(0.014)	(0.035)
First Stage F-Stat		351.11	21.12
First Stage Partial R^2		0.31	0.05
Observations	26,884	26,884	19,116
Specification	OLS	IV	IV
Instrument		Sensitivity	Saiz Elast
Num. CBSAs	380	380	270
Yr-Qtr FE			✓
Region, Yr-Qtr FE	✓	✓	
CBSA FE	✓	✓	✓

Notes: Replication of Guren et al. (2021), table 1 using the equation $\Delta y_{i,r,t} = \psi_i + \xi_{r,t} + \beta \Delta p_{i,r,t} + \Gamma X_{i,r,t} + \epsilon_{i,r,t}$. $\Delta y_{i,r,t}$ is the log annual change in quarterly retail employment per capita (a consumption proxy in year-over-year first-difference form) for CBSA i in census region r at time t. $\Delta p_{i,r,t}$ is the log annual change in quarterly house prices for CBSA i. ψ_i, $\xi_{r,t}$, and $X_{i,r,t}$ represent CBSA fixed effects, census region × time fixed effects, and other controls, such as industry shares, respectively. See Guren et al. (2021) for a full list of controls. Robust standard errors clustered by time and CBSA are in parentheses. One, two, or three asterisks represent statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A5: 2SLS Housing Wealth Elasticity Estimates as in Guren et al. (2021)

	YoY Log Diff in Retail Emp Per Capita				
	(1)	(2)	(3)	(4)	(5)
Panel A: 1978–2017					
YoY Log Diff in House Prices	0.085***	0.082	0.057***	0.289	0.079***
	(0.007)	(0.054)	(0.017)	(0.296)	(0.016)
First Stage F-Stat		14.37	235.16	1.12	297.48
First Stage Partial R^2		0.02	0.17	0.00	0.17
Panel B: 1990–2017					
YoY Log Diff in House Prices	0.081***	0.136***	0.072***	0.583	0.082***
	(0.008)	(0.040)	(0.015)	(0.703)	(0.016)
First Stage F-Stat		16.34	434.18	0.67	236.18
First Stage Partial R^2		0.04	0.28	0.00	0.21
Panel C: 2000–2017					
YoY Log Diff in House Prices	0.067***	0.121***	0.055***	0.427	0.070***
	(0.008)	(0.037)	(0.014)	(0.761)	(0.018)
First Stage F-Stat		16.26	347.48	0.32	171.03
First Stage Partial R^2		0.05	0.31	0.00	0.20
Specification	OLS	IV	IV	IV	IV
Instrument		Saiz	Sensitivity	Baum-Snow	LU-ML
Num. CBSAs	376	270	376	311	376
CBSA FE	✓	✓	✓	✓	✓
Region × Date FE	✓	✓	✓	✓	✓

Notes: Columns (1) to (3) replicate Guren et al. (2021) using the equation $\Delta y_{i,r,t} = \psi_i + \xi_{r,t} + \beta \Delta p_{i,r,t} + \Gamma X_{i,r,t} + \epsilon_{i,r,t}$. $\Delta y_{i,r,t}$ is the log annual change in quarterly retail employment per capita (a consumption proxy in year-over-year first-difference form) for CBSA i in census region r at time t. $\Delta p_{i,r,t}$ is the log annual change in quarterly house prices for CBSA i. ψ_i, $\xi_{r,t}$, and $X_{i,r,t}$ represent CBSA fixed effects, census region × time fixed effects, and other controls, such as industry shares, respectively. See Guren et al. (2021) for a full list of controls. Column (1) employs OLS, while columns (2) and (3) use the Saiz Elasticity and Sensitivity instruments, respectively. Column (4) uses the elasticity proxy from Baum-Snow and Han (2024) and column (5) employs the LU-ML IV. Robust standard errors clustered by time and CBSA are in parentheses. One, two, or three asterisks represent statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A6: Long-Difference Regressions of House Price Growth on BSH, Elasticity, Saiz Elasticity, and LU-ML Price Pressure

	YoY Log Diff in House Prices			
	(1) 2000-2007	(2) 2007-2011	(3) 2012-2019	(4) 2019-2024
Panel A: BSH Elasticity				
BSH Elasticity	0.164	−0.061	−0.179**	0.115***
	(0.103)	(0.085)	(0.077)	(0.035)
First Stage F-Stat	2.52	0.51	5.38	10.47
First Stage Partial R^2	0.01	0.00	0.02	0.04
Number of CBSAs	314	314	314	314
Panel B: Saiz Elasticity				
Saiz Elasticity	−0.081***	0.063***	−0.053***	−0.007
	(0.014)	(0.012)	(0.009)	(0.005)
First Stage F-Stat	31.62	28.80	32.91	2.70
First Stage Partial R^2	0.26	0.19	0.17	0.01
Number of CBSAs	251	251	251	251
Panel C: LU-ML Price Pressure				
LU-ML Price Pressure	1.088***	0.984***	1.015***	1.054***
	(0.055)	(0.085)	(0.078)	(0.133)
First Stage F-Stat	388.91	133.25	167.40	63.17
First Stage Partial R^2	0.54	0.38	0.39	0.19
Number of CBSAs	380	380	380	380

Notes: Regressions of the long difference in Freddie Mac CBSA house prices on Saiz elasticity, BSH Elasticity, or the LU-ML Price Pressure Index. We compute the LU-ML Price Pressure Index by letting the target variable in our LU-ML approach be the log difference in Freddie Mac house prices. Estimates use the equation $\Delta \log HP_i = \beta_0 + \beta_1 \cdot \mathrm{Z}_i + \varepsilon_i$. $\Delta \log HP_i$ is the long difference in house prices for CBSA i and Z_i is the Saiz Elasticity proxy (panel A), the BSH elasticity measure (panel B), or LU-ML Price Pressure (panel C). Robust standard errors clustered by commuting zone (2020 definitions). One, two, or three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A7: Panel Regressions of House Price Growth on BSH Elasticity, Saiz Elasticity, and LU-ML Price Pressure Index

	YoY Log Diff in House Prices			
	(1) 1976-2024	(2) 1976-1999	(3) 2000-2011	(4) 2012-2024
Panel A: BSH Elasticity				
BSH Elasticity	0.017	−0.293**	0.185	0.112
	(0.117)	(0.130)	(0.171)	(0.086)
First Stage F-Stat	0.02	5.07	1.17	1.72
First Stage Partial R^2	0.00	0.00	0.00	0.00
Number of CBSAs	309	309	309	309
Panel B: Saiz Elasticity				
Saiz Elasticity	−0.071***	−0.042**	−0.100***	−0.020
	(0.021)	(0.021)	(0.029)	(0.013)
First Stage F-Stat	10.93	3.99	11.65	2.52
First Stage Partial R^2	0.02	0.00	0.05	0.00
Number of CBSAs	249	249	249	249
Panel C: LU-ML Price Pressure				
LU-ML Price Pressure	0.607***	0.503***	0.751***	0.506***
	(0.046)	(0.050)	(0.065)	(0.064)
First Stage F-Stat	171.00	101.14	131.89	63.27
First Stage Partial R^2	0.14	0.09	0.25	0.09
Number of CBSAs	380	380	380	380
CBSA FE	✓	✓	✓	✓
Division × Date FE	✓	✓	✓	✓

Notes: Regressions of the year-over-year (YoY) log first-difference in Freddie Mac CBSA house prices on Saiz elasticity, BSH Elasticity, and the LU-ML Price Pressure Index. We compute the LU-ML Price Pressure Index by letting the target variable in our LU-ML approach be the YoY log difference in Freddie Mac house prices. Estimates use the equation: $\Delta \log HP_{i,d,t} = \psi_i + \gamma_{d,t} + \beta \cdot SC_{i,d,t} + \varepsilon_{i,d,t}$. $\Delta \log HP_{i,d,t}$ is the YoY log first-difference in house prices for CBSA i in Census Division d at month t, ψ_i are CBSA fixed effects, and $\gamma_{d,t}$ are division-by-date fixed effects. There are nine Census divisions. $SC_{i,d,t}$ is the Saiz Elasticity proxy (panel A), the BSH Elasticity measure (panel B), or the LU-ML Price Pressure Index (panel C) for CBSA i in Census division d at time t. To use the Saiz or BSH elasticity proxies within a panel framework, we multiply it by the year-over-year log change in the national house price as in Guren et al. (2021). If we multiply Saiz or BSH elasticity measure by annual changes in the real national mortgage rate, as in Chaney et al. (2012), the results are similar. Robust standard errors clustered by commuting zone (2020 definitions) are in parentheses. One, two, or three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A8: Supply Constraint Predictive Power Summary Statistics, 2000-2011

Geog Level	HP Dataset	Baum-Snow and Han	Saiz	LU-ML Price Pressure
CBSA	FHFA	$R^2 = 0.00$, N = 332	$R^2 = 0.05$, N = 256	$R^2 = 0.19$, N = 910
CBSA	Freddie Mac	$R^2 = 0.00$, N = 309	$R^2 = 0.05$, N = 249	$R^2 = 0.25$, N = 380
CBSA	Zillow	$R^2 = 0.00$, N = 176	$R^2 = 0.04$, N = 147	$R^2 = 0.24$, N = 302
County	FHFA	$R^2 = 0.00$, N = 711	$R^2 = 0.05$, N = 1582	$R^2 = 0.31$, N = 2395
County	Zillow	$R^2 = 0.00$, N = 355	$R^2 = 0.07$, N = 583	$R^2 = 0.38$, N = 735
Zip3	FHFA	$R^2 = 0.00$, N = 619	$R^2 = 0.07$, N = 713	$R^2 = 0.19$, N = 860
Zip3	FHFA	$R^2 = 0.00$, N = 554	$R^2 = 0.08$, N = 589	$R^2 = 0.20$, N = 637
Census Tract	FHFA	$R^2 = 0.00$, N = 40374	$R^2 = 0.05$, N = 39215	$R^2 = 0.47$, N = 48636
Zip5	FHFA	$R^2 = 0.01$, N = 11635	$R^2 = 0.06$, N = 11357	$R^2 = 0.40$, N = 14814
Zip5	Zillow	$R^2 = 0.01$, N = 9682	$R^2 = 0.06$, N = 9367	$R^2 = 0.53$, N = 11705

Notes: Partial R^2 statistics and cross-section geography counts (N) across housing panel datasets using the Saiz (2010) or Baum-Snow and Han (BSH, 2024) elasticity proxies versus the LU-ML Price Pressure Index. Housing datasets include those from Freddie Mac, the FHFA, and Zillow, aggregated to zip codes, census tracts, three-digit zip codes, counties, or CBSAs. We compute the LU-ML Price Pressure Index by letting the target variable in our LU-ML approach be the YoY log difference in house prices, separately, for each of these datasets. Estimates use the equation: $\Delta \ln HP_{i,d,t} = \psi_i + \gamma_{d,t} + \beta \cdot Z_{i,d,t} + \varepsilon_{i,d,t}$. $\Delta \ln HP_{i,d,t}$ is the YoY log first-difference in house prices for geographic unit i (e.g., a CBSA or zip code) in Census Division d at month t, ψ_i are geographic unit fixed effects, and $\gamma_{d,t}$ are division-by-date fixed effects. There are nine Census divisions. $Z_{i,d,t}$ is the Saiz Elasticity proxy, the BSH Elasticity measure, or the LU-ML Price Pressure Index for geographic unit i in Census division d at time t. To use the Saiz or BSH elasticity proxies within a panel framework, we multiply them by the year-over-year log change in the national house price as in Guren et al. (2021).

Table A9: Supply Constraint Predictive Power Summary Statistics, 2000-2024

Geog Level	HP Dataset	Baum-Snow and Han	Saiz	LU-ML Price Pressure
CBSA	FHFA	$R^2 = 0.00$, N = 332	$R^2 = 0.03$, N = 256	$R^2 = 0.13$, N = 908
CBSA	Freddie Mac	$R^2 = 0.00$, N = 309	$R^2 = 0.03$, N = 249	$R^2 = 0.20$, N = 380
CBSA	Zillow	$R^2 = 0.00$, N = 172	$R^2 = 0.03$, N = 144	$R^2 = 0.17$, N = 292
County	FHFA	$R^2 = 0.00$, N = 710	$R^2 = 0.02$, N = 1562	$R^2 = 0.20$, N = 2334
County	Zillow	$R^2 = 0.00$, N = 348	$R^2 = 0.03$, N = 561	$R^2 = 0.32$, N = 706
Zip3	FHFA	$R^2 = 0.00$, N = 618	$R^2 = 0.03$, N = 712	$R^2 = 0.15$, N = 857
Zip3	FHFA	$R^2 = 0.00$, N = 554	$R^2 = 0.03$, N = 589	$R^2 = 0.16$, N = 637
Census Tract	FHFA	$R^2 = 0.00$, N = 29638	$R^2 = 0.02$, N = 28871	$R^2 = 0.35$, N = 35554
Zip5	FHFA	$R^2 = 0.00$, N = 10454	$R^2 = 0.02$, N = 10195	$R^2 = 0.30$, N = 12979
Zip5	Zillow	$R^2 = 0.01$, N = 9567	$R^2 = 0.03$, N = 9270	$R^2 = 0.50$, N = 11535

Notes: See the notes for table A8.

B Appendix: Income Growth and Housing Markets by Supply Elasticities

Table A10: Impact of Income Growth on Housing Markets by Supply Proxy

	Supply Constraint Measure:		
	BSH	Saiz	LU-ML
	(1)	(2)	(3)
Panel A: LHS = House Price Growth			
Total Income Growth	0.828***	0.644***	0.739***
	(0.067)	(0.096)	(0.064)
LessConstrained	0.044	−0.225**	0.293***
	(0.081)	(0.097)	(0.107)
Total Income Growth × *LessConstrained*	−0.086	0.123	−0.252**
	(0.102)	(0.120)	(0.126)
Number of CBSAs	314	251	380
Panel B: LHS = Housing Unit Growth			
Per Capita Income Growth	0.303***	0.153	−0.058
	(0.117)	(0.178)	(0.078)
LessConstrained	0.011	−0.207	−0.129
	(0.108)	(0.143)	(0.085)
Per Capita Income Growth × *LessConstrained*	−0.023	0.262	0.322**
	(0.163)	(0.222)	(0.128)
Number of CBSAs	332	256	919

Notes: CBSA-level regressions from 2000-2020 using the equation:

$$\Delta \log(Y_i) = \alpha + \beta_1 \Delta \log(Income_i) + \beta_2 LessConstrained_i + \beta_3[\Delta \log(Income_i) \times LessConstrained_i] + \epsilon_i$$

Y_i is the Freddie Mac house price for CBSA i in panel A and the number of housing units in Panel B (tabulated to CBSAs from Census block groups using NHGIS data). Differences are taken from 2000 to 2020. In panel A, income growth uses total income from the BEA, while panel B uses BEA per capita income. $LessConstrained_i$ is an indicator equal to 1 for CBSAs with above average supply constraints. Column (1) uses the Baum-Snow and Han (2024, BSH) elasticity proxy, column (2) employs the Saiz (2010) elasticity variable, and column (3) uses the LU-ML Supply Index. We compute the LU-ML Supply Index by letting the target variable in our LU-ML approach be the log change in CBSA-level housing units. The regression intercept is not shown. All data are aggregated to 2020 CBSA delineations. Robust standard errors are clustered at the 2020 commuting-zone level. One, two, or three asterisks represent statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A11: Estimated Income Elasticities of House Prices – More versus Less Constrained

Estimation Type	Geographic Level	HP Dataset	SC Measure	Constrained Elasticity	Elasticity Diff (Less − More)
2000-2020 Long Difference	Major Metros	Freddie Mac	BSH	0.828 (0.067); $p = 0.000$	-0.086 (0.102); $p = 0.399$
2000-2020 Long Difference	Major Metros	Freddie Mac	Saiz	0.644 (0.096); $p = 0.000$	0.123 (0.120); $p = 0.307$
2000-2020 Long Difference	Major Metros	Freddie Mac	LU-ML	0.739 (0.064); $p = 0.000$	-0.252 (0.126); $p = 0.045$
2000-2020 Long Difference	Major Metros	FHFA	BSH	0.796 (0.064); $p = 0.000$	-0.104 (0.094); $p = 0.269$
2000-2020 Long Difference	Major Metros	FHFA	Saiz	0.615 (0.089); $p = 0.000$	0.104 (0.110); $p = 0.347$
2000-2020 Long Difference	Major Metros	FHFA	LU-ML	0.715 (0.067); $p = 0.000$	-0.264 (0.124); $p = 0.034$
2000-2020 Long Difference	Major Metros	Zillow	BSH	0.624 (0.091); $p = 0.000$	-0.128 (0.133); $p = 0.338$
2000-2020 Long Difference	Major Metros	Zillow	Saiz	0.475 (0.107); $p = 0.000$	0.073 (0.156); $p = 0.638$
2000-2020 Long Difference	Major Metros	Zillow	LU-ML	0.617 (0.101); $p = 0.000$	-0.257 (0.141); $p = 0.069$
2000-2020 Long Difference	All CBSAs	FHFA	LU-ML	0.721 (0.047); $p = 0.000$	-0.163 (0.082); $p = 0.048$
2000-2020 Long Difference	All Counties	FHFA	LU-ML	0.502 (0.053); $p = 0.000$	-0.106 (0.058); $p = 0.067$
1990-2020 Decadal Panel	All CBSAs	FHFA	LU-ML	0.827 (0.068); $p = 0.000$	-0.221 (0.065); $p = 0.001$
1990-2020 Decadal Panel	All Counties	FHFA	LU-ML	0.677 (0.066); $p = 0.000$	-0.204 (0.057); $p = 0.000$

Notes: The table shows the estimated impact of income growth on house price growth for *Less Constrained* versus *More Constrained* CBSAs, by supply constraint proxy. *Less Constrained* CBSAs have looser-than-average supply constraints. Supply proxies are from Baum-Snow and Han (2024, BSH), Saiz (2010), and the LU-ML Supply Index. We compute LU-ML Supply Index using the log change in housing units from 2000-2020 as the target variable in our LU-ML approach. Long-difference estimates from 2000-2020 using the following regression:

$$\Delta \log(HP_i) = \alpha + \beta_1 \Delta \log(Income_i) + \beta_2 LessConstrained_i + \beta_3[\Delta \log(Income_i) \times LessConstrained_i] + \epsilon_i$$

We run separate regressions using each supply constraint proxy (BSH, Saiz, LU-ML). Decadal panel regression from 1990-2020 use the FHFA house price dataset with all available U.S. data in this regression:

$$\Delta \log(HP_{it}) = \alpha_i + \eta_t + \beta_1 \Delta \log(Income_{it}) + \beta_3[\Delta \log(Income_{it}) \times LessConstrained_i] + \epsilon_{it}$$

α_i are CBSA/county fixed effects η_t are decadal-time fixed effects. All data are aggregated to 2020 CBSA or county delineations. Robust standard errors are clustered at the 2020 commuting-zone level.

Table A12: Average Percentage Change in House Prices by Davidoff (2013) Market Category

Market	2000-2015	2000-2020	2010-2020
Sand and Coastal	63.882	122.384	60.430
Sand Only	48.869	104.722	59.556
Coastal Only	43.846	78.248	26.282
Other	36.888	65.781	26.010

Notes: The percentage change in CBSA house price growth, using FHFA annual house prices, by Davidoff (2013) market category. Davidoff defines sand states as Arizona, California, Florida, and Nevada. Coastal markets are defined as those CBSAs that intersect with the U.S. coastline.

Table A13: Average Percentage Change in House Prices by Davidoff (2013) Market Category Including Texas

Market	2000-2015	2000-2020	2010-2020
Sand + TX and Coastal	62.887	116.766	55.614
Sand + TX Only	56.540	103.385	48.524
Coastal Only	42.986	77.660	26.007
Other	34.555	62.530	24.828

Notes: The percentage change in CBSA house price growth, using FHFA annual house prices, by Davidoff (2013) market category including Texas as a sand state. Davidoff defines sand states as Arizona, California, Florida, and Nevada. This table also includes Texas. Coastal markets are defined as those CBSAs that intersect with the U.S. coastline.

Table A14: Income Elasticity of House Prices – Interactions with Continuous LU-ML Supply Index (1990-2020 County Decadal Panel)

	Dependent variable:			
	$\Delta \log$ HPI			
	(1)	(2)	(3)	(4)
Total Income Growth	0.456^{***}	0.371^{***}	0.390^{***}	0.570^{***}
	(0.028)	(0.028)	(0.031)	(0.045)
LU-ML Supply Index (Standardized)	0.029^{***}	0.025^{***}	0.034^{***}	
	(0.008)	(0.007)	(0.009)	
Total Income Growth ×	-0.062^{***}	-0.057^{***}	-0.072^{***}	-0.079^{***}
LU-ML Supply Index (Standardized)	(0.019)	(0.018)	(0.019)	(0.026)
Decadal-Time FE	✓	✓	✓	✓
State FE		✓		
Commuting Zone FE			✓	
County FE				✓

Notes: This table reports the estimated elasticity of house prices with respect to income, interacting income growth with the continuous LU-ML Supply Index. We compute LU-ML Supply Index using the log change in housing units from 2000-2020 as the target variable in our LU-ML approach. The LU-ML Supply Index is standardized to have mean zero and unit variance. The coefficient on $\Delta \log(\textit{Total Income})$ (β_1) represents the estimated elasticity for a location with *average* supply constraints. The interaction coefficient (β_3) estimates how the elasticity changes with a one standard deviation increase in the supply index (indicating less constrained supply). The estimates are based on the following specification, where α_i and η_t represent geographic and time fixed effects, respectively:

$$\Delta \log(HP_{it}) = \alpha_i + \eta_t + \beta_1 \Delta \log(Income_{it}) + \beta_2 LU\text{-}ML\ Index_i + \beta_3[\Delta \log(Income_{it}) \times LU\text{-}ML\ Index_i] + \epsilon_{it}$$

The main effect of the time-invariant supply index is absorbed by the county fixed effects in Column (4). Robust standard errors are clustered at the 2020 commuting-zone level. One, two, and three asterisks represents statistical significance at the 10, 5, and 1 percent levels, respectively.