Online Appendix to: "The Effect of Land Supply for New Homes on Residential Investment and House Prices"

Justin Katz and Paul Willen

April 2026

Table of Contents

A CoreLogic data cleaning.

We start with the Boston Fed's maintained version of the CoreLogic Historical Property Basic dataset for Massachusetts, Connecticut, Rhode Island, Vermont, New Hampshire, Maine, Virginia, North Carolina, South Carolina, Georgia, Tennessee, Mississippi, Louisiana, and Arkansas. We follow steps similar to our processing of the Warren Group data.

Data cleaning. We drop rows with missing parcel identifiers, acreage over 10,000, and deduplicate on parcel ID, assessed year, tax year, acres, and land use codes. This forms the dataset used to compare coverage across time in southern New England states.

2019 sample selection. We then filter to parcels with nonmissing 2019 data. Assessors update parcel information as of the year given in the `assessed_year` field. Since we want information from 2019, we select one observation per parcel as follows. we keep the first tax year observation with assessed year of 2019. If a parcel was not assessed in 2019, we keep the tax year observation with tax year of 2019. If a parcel was not assessed in 2019 and has no tax year 2019 observation, we keep the earliest post-2019 tax year observation with a pre-2019 assessed year.

Land use classifications. We do not fully classify all land use codes as with the Warren Group data, because we only use CoreLogic for select analyses. We do, however, classify parcels as residential vacant or not. We do so with CoreLogic's harmonized `land_use_code` field. We classify a parcel as residential vacant if it falls into one of the following categories:

- 450: Multi-family acreage
- 452: Multi-family lot
- 453: Residential open space
- 454: Vacant mobile home
- 460: Residential acreage
- 465: Residential lot

Parcels with these land use codes have de-minimus assessed structure value, consistent with reflecting vacant land.

B Model details.

B.1 Derivation of equation (5).

From the FOC $C'_{mt}(H^*) = P_{mt}$ and the parameteriation of $C_{mt}(\cdot)$:

$$c_{mt} \cdot H^{\frac{1}{\alpha-1}} = P_{mt} \implies H^* = c_{mt}^{1-\alpha} \cdot P^{\alpha-1} \tag{B.1}$$

Hence:

$$\begin{aligned} \pi_{mt}(P) &= \alpha^{-1} \cdot c_{mt}^{1-\alpha} \cdot P^{\alpha} && \text{(B.2)} \\ &= \pi_{0mt} \cdot P^{\alpha} && \text{(B.3)} \end{aligned}$$

where $\pi_{0mt} := \alpha^{-1} \cdot c_{mt}^{1-\alpha}$. Multiplying variable profits from floorspace per acre by parcel acreage, and subtracting parameterized fixed costs gives equation (5).

B.2 Calibration details.

This subsection gives additional details on model calibration. Table TK summarizes results.

House prices and quantities. We normalize 2018 house prices to 1. We estimate real price growth from 2018-2019 of 2.9% using the FHFA repeat-sales price index for New England, deflated with the CPI-U. We normalize the total mass of buildable land to 1. We calibrate the 2018 housing stock to 4.5, which matches the acreage-weighted ratio of residential developed land to residential vacant land in 2018 in the estimation sample for Table 5. We define housing markets as zip codes, and calculate the average parcel size in acres of residential vacant parcels to form the baseline distribution A_t.

Developers. We calculate N_m as the ratio of total residential vacant acreage in zipcode m and average parcel size in zipcode m, divided by the total residential acreage across all markets m. This ensures that $\sum_m N_m \cdot a_{mt} = 1$. We take $\rho = 0.60$ and $\alpha = 3.65$ from column (3) of Table 5. We calculate $\ln(\pi_{0mt}/f_{mt})$ as the residuals in equation (9), winsorized at the 2.5th and 97.5th percentiles. By construction, residuals are mean-zero across all markets; we shift the mean by a factor $\beta = -3.33$ selected so that the modeled 2019 development rate matches the average empirical development rate in the estimation sample for Table 5.

Housing demand. We calibrate the price elasticity of housing demand as $\eta_1 = 1.0$, roughly the average demand elasticity estimated in Chodorow-Reich *et al.* (2024) (see Table 2). We set the demand intercept $\eta_0 = -2.54$ so that the modeled 2018 equilibrium price equals 1.

Table B.1: Calibrated Parameters

Variable	Symbol	Value	Source
Prices and quantities			
2018 house price	P_{2018}	1	Normalization
Real price growth, 2018–2019	ΔP_{2019}	2.9%	FHFA repeat-sales index, CPI-U
Total mass of buildable land, 2019	$\sum_m N_m a_{m,2019}$	1	Normalization
2018 housing stock	Q_{2018}	4.5	Warren Group
Parcel size distribution	$\{a_{m,2019}\}_m$	3.2 (2.9)	Warren Group
Developer mass	$\{N_m\}_m$	4.9×10^{-4} (6.6×10^{-4})	Warren Group
Developers			
Supply elasticity	ρ	0.60	Table 5, col. (3)
Supply curvature	α	3.65	Table 5, col. (3)
Log profit-cost residual	$\{\ln(\pi_{0m,2019}/f_{m,2019})\}_m$	-2.6 (2.8)	Warren Group
Mean-shift factor	β	−3.33	Warren Group
Housing demand			
Price elasticity of housing demand	η_1	1.0	Calibrated
Demand intercept	η_0	−2.54	Calibrated

The table summarizes calibrated parameters. For rows with multiple parameter values (e.g., the parcel size distribution), the Value column reports the mean and standard deviation across the parcels in the estimation sample.

B.3 Productivity calculations.

We define productivity as total fixed and variable costs per unit of floorspace.

Variable costs per developed acre are:

$$c_{mt} \cdot \frac{\alpha - 1}{\alpha} \cdot (H^*)^{\frac{\alpha}{\alpha-1}} = c_{mt}^{1-\alpha} \cdot \frac{\alpha - 1}{\alpha} \cdot P^{\alpha} \tag{B.4}$$

Fixed costs per developed acre are:

$$f_{mt} \cdot a_{mt}^{-\rho} \cdot \exp\left(\epsilon_{jmt}\right) \tag{B.5}$$

Taking expectations over j yields:

$$f_{mt} \cdot a_{mt}^{-\rho} \cdot E\left[\exp\left(\epsilon_{jmt}\right) | \epsilon_{jmt} \leq \ln(\pi_{0mt}/f_{mt}) + \alpha \ln P_{mt}\right] \tag{B.6}$$

Note that for x distributed logistic with cdf F:

$$E[\exp(x)|x \leq c] = F(c)^{-1} \cdot \int_{-\infty}^{c} \exp(x) dF(x) \tag{B.7}$$

$$= F(c)^{-1} \cdot \int_{-\infty}^{c} \frac{\exp(x)\exp(-x)}{(1+\exp(-x))^2} dx \tag{B.8}$$

$$= F(c)^{-1} \cdot \int_{-\infty}^{c} \frac{1}{(1+\exp(-x))^2} dx \tag{B.9}$$

Substituting $u = \frac{1}{1+\exp(-x)}$, meaning $\frac{du}{u(1-u)} = dx$:

$$\int_{-\infty}^{c} \frac{1}{(1+\exp(-x))^2} dx = \int_{u(-\infty)}^{u(c)} \frac{u}{1-u} du \tag{B.10}$$

$$= [-u - \ln(1-u)]_0^{u(c)} \tag{B.11}$$

$$= -u(c) - \ln(1-u(c)) \tag{B.12}$$

$$= \ln(1+\exp(c)) - F(c) \tag{B.13}$$

Hence:

$$E\left[\exp\left(\epsilon_{jmt}\right) | \epsilon_{jmt} \leq \ln(\pi_{0mt}/f_{mt}) + \alpha \ln P_{mt}\right] = \frac{\ln\left(1 + \pi_{0mt}/f_{mt} + P_{mt}^{\alpha}\right)}{q_{mt}} - 1 \tag{B.14}$$

Putting everything together, productivity is:

$$\frac{c_{mt}^{1-\alpha} \cdot P^{\alpha-1}}{f_{mt} a_{mt}^{-\rho} E[\epsilon_{jmt}|\cdot] + \frac{\alpha-1}{\alpha} c_{mt}^{1-\alpha} P^{\alpha}} = \frac{P^{\alpha-1}}{(\alpha\pi_{0mt}/f_{mt})^{-1} a_{mt}^{-\rho} E[\epsilon_{jmt}|\cdot] + \frac{\alpha-1}{\alpha} P^{\alpha}} \tag{B.15}$$

We have calibrated $\alpha, P, a_{mt}, q_{mt}$, and π_{0mt}/f_{mt}, so we can calculate productivity using the formula in equation (B.15). As explained in the main text, all of our productivity calculations hold P and q fixed at their 2018 baseline levels.

C Additional tables.

Table C.2: Price growth by buildable land distribution – county fixed effects.

	% Δ Price, 2013-2019		
	(1)	(2)	(3)
Variables			
nthou_hhlds_2013	0.0037**		0.0027**
	(0.0014)		(0.0011)
Ln(Income), 2013	-0.1122***		-0.1085***
	(0.0206)		(0.0208)
Δ Ln(Income), 2013-2019	0.0239		0.0221
	(0.0346)		(0.0327)
Price / rent, 2013	-0.0007		-0.0007
	(0.0005)		(0.0005)
Acres, p90 residential vacant		-0.0033***	-0.0019**
		(0.0010)	(0.0008)
Fixed-effects			
County	Yes	Yes	Yes
Observations	728	728	728
R^2	0.78254	0.72642	0.78861
Within R^2	0.26498	0.07531	0.28550

SEs clustered by county
****: 0.01, **: 0.05, *: 0.1*

Source: Warren Group. This table replicates Table 4, except includes county, rather than state, fixed effects. See the notes to Table 4 for details.

Table C.3: CoreLogic coverage by state.

State	**Coverage**
Southern New England	
CT	0.732
MA	0.694
RI	0.360
Northern New England	
ME	0.200
NH	0.636
VT	0.875
The South	
AR	0.834
GA	0.832
LA	0.381
MS	0.688
NC	0.744
SC	0.892
TN	0.849
VA	0.835

Source: CoreLogic. This table displays the fraction of total noncoastal area in each state covered by parcels in CoreLogic data in 2019.

References

CHODOROW-REICH, G., GUREN, A. M. and MCQUADE, T. J. (2024). The 2000s housing cycle with 2020 hindsight: A neo-kindlebergerian view. *Review of Economic Studies*, **91** (2), 785–816.