Beyond the Home–Work Separation: A RAG-Based Analysis of U.S. Municipal Ordinances and Remote Work

Keywords: work from home; zoning; land-use regulation; retrieval-augmented generation; urban form.

Extended Abstract

Work from home (WFH) refers to paid work performed primarily from one's residence. Since the COVID-19 shock, WFH has become a durable feature of the U.S. labor market [1]. According to American Community Survey (ACS) five-year estimates ending in 2019 and 2024, the WFH share of workers rises from about 5% before the pandemic to 16% afterward, with highly uneven growth across places (Fig. A1). Beyond its scale, WFH dissolves the premise on which a century of Euclidean zoning rests—the spatial *separation of home from work*—and reshapes spatial demand, as commuting falls and activity shifts from central offices toward homes and nearby neighborhoods [2, 3]. Land-use scholars therefore argue that zoning, organized around commuting, must now actively adapt to remote work rather than merely tolerate it [4].

However, systematic quantitative evidence remains limited on which specific zoning provisions are most relevant to local adaptation under WFH. A central obstacle is that zoning authority is highly decentralized, while zoning codes are lengthy, heterogeneous, and difficult to collect at scale. Recent advances in NLP and LLMs make large-scale zoning measurement increasingly feasible, but existing efforts remain incomplete in coverage and mainly extract residential restrictiveness or general land-use variables rather than WFH-relevant provisions [6–9]. At the same time, WFH studies explain adoption by worker and industry composition and document shifts from central office districts toward homes and neighborhoods [5, 2, 3]. The institutional role of zoning in this transition therefore remains undermeasured.

To address this gap, we collect what is, to our knowledge, the largest national zoning corpus from AmLegal, Municode, and ordinance.com. We preserve the hierarchy of the original legal text, geomap jurisdictions, extract zoning sections, chunk the text, and use Qwen3-8B embeddings to build a vector database for 8,225 U.S. jurisdictions (≈1.7M text chunks). The outcome variable is the percentage-point change in the American Community Survey (ACS) five-year WFH share between the estimates ending in 2019 and 2024 (ΔWFH). Jurisdictions are matched to ACS places using official U.S. Census Bureau GEOIDs, and places with at least 1,000 workers in both years are kept, yielding a main sample of 4,650 places (Table A1).

We then use an expert-guided question set of 12 regulatory dimensions that capture zoning margins relevant to WFH adaptability (Table A2), which cover direct home-occupation rules and indirect provisions for neighborhood-scale flexibility, including mixed-use/live-work, parking flexibility, and accessory dwelling units (ADUs). For each jurisdiction, hybrid lexical/semantic search plus cross-encoder reranking retrieves relevant passages; an LLM ensemble (Qwen3.6-35B, DeepSeek-v4-Flash, and Gemma4-31B) codes each dimension with source-text evidence, cross-model checks, and manual validation (Fig. A2; Table A1, Panel C).

Empirically, the analysis proceeds in two stages. The regression stage estimates per-variable complete-case models of ΔWFH on standardized RAG-coded zoning measures, run separately for direct and indirect dimensions; specifications include 2019 ACS controls, are estimated with and without state fixed effects, and cluster standard errors by state (Table A3). The prediction stage uses held-out tests to compare ACS-only models with models that add the 12 RAG dimensions and/or PCA-reduced full-code document embeddings, using OLS, Lasso, histogram gradient boosting, and XGBoost (Table A4).

Table A3 first shows that direct home-occupation rules do not predict WFH growth. With ACS controls and state fixed effects, all seven coefficients are small and insignificant. This is counterintuitive but explainable. Home-occupation rules govern visible business activity: employees, clients, signs, deliveries, storage, floor area, noise, and traffic. ACS WFH mainly captures salaried employees working from home. Such remote work is hard to observe and rarely triggers complaint-driven enforcement, so de jure limits are not binding. In addition, most places already permit some home occupation, leaving little relevant cross-place variation beyond fine rules such as employee caps or visit restrictions. The indirect neighborhood-adaptation margin is different: it governs whether residential areas can absorb new daytime activity, service demand, and flexible space needs. With ACS controls and state fixed effects, one-SD increases in ADU permission, mixed-use/live-work, and parking flexibility are associated with +0.222, +0.250, and +0.230 pp higher WFH growth, respectively; apartments above commercial and no off-street parking minimums are near zero. These coefficients suggest that WFH growth aligns more with neighborhood flexibility—housing adaptability, mixed local uses, and relaxed parking constraints—than with formal home-business permission. Prediction results provide a complementary text-as-data check (Table A4). Embedding-only models explain a nontrivial share of held-out variation, though less than the ACS baseline. Adding RAG measures and PCA-100 embeddings still yields modest incremental gains after ACS controls and state fixed effects, suggesting that zoning text contains additional institutional signal while part of that signal is spatially structured and absorbed by state fixed effects.

Our results locate where remote-work growth co-varies with land-use regulation: the association concentrates in ADU permission, mixed-use/live-work, and parking flexibility rather than direct home–work legality. These are descriptive conditional associations, not causal effects; their modest magnitudes, residence-side outcome, and modest predictive gains after ACS controls and state fixed effects are consistent with sorting of remote-capable workers into flexibility-friendly places. The ordinance snapshot is from 2025, but zoning rules are often durable, so the cross-section is best read as institutional context rather than observed policy change. Future work will move to higher-frequency, workplace-linked WFH and home-business measures to test whether the home-occupation null persists; collect multi-vintage ordinances and datable local reforms for staggered difference-in-differences and border-discontinuity designs; and further validate the RAG instrument with expanded cross-model and human audits.

References

1. Barrero, J. M., Bloom, N., & Davis, S. J. (2023). The evolution of work from home. *Journal of Economic Perspectives, 37*(4), 23–50.
2. Ramani, A., Alcedo, J., & Bloom, N. (2024). How working from home reshapes cities. *Proceedings of the National Academy of Sciences, 121*(45), e2408930121.
3. Gupta, A., Mittal, V., & Van Nieuwerburgh, S. (2026). Work from home and the office real estate apocalypse. *American Economic Review, 116*(2), 674–709.
4. Stern, S. M. (2022). UNTRANSIT: Remote work and the transformation of zoning. *Stanford Law & Policy Review, 33*(1), 79–132.
5. Dingel, J. I., & Neiman, B. (2020). How many jobs can be done at home? *Journal of Public Economics, 189*, 104235.
6. Gyourko, J., Hartley, J. S., & Krimmel, J. (2021). The local residential land use regulatory environment across U.S. housing markets: Evidence from a new Wharton index. *Journal of Urban Economics, 124*, 103337.
7. Mleczko, M., & Desmond, M. (2023). Using natural language processing to construct a National Zoning and Land Use Database. *Urban Studies, 60*(13), 2564–2584.
8. Salazar-Miranda, A., & Talen, E. (2025). An AI-based analysis of zoning reforms in US cities. *Nature Cities, 2*, 304–315.
9. Bartik, A. W., Gupta, A., & Milo, D. (2025, August 19). The costs of housing regulation: Evidence from generative regulatory measurement. *SSRN*.

Appendix: Descriptive Statistics and Figures

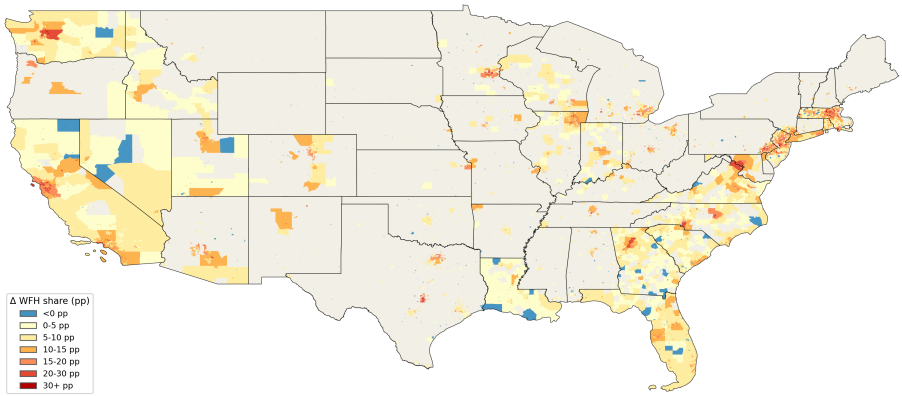

Fig. A1: Spatial variation in the change in ACS five-year WFH share among matched zoning units. The map reports percentage-point changes between ACS estimates ending in 2019 (2015–2019) and 2024 (2020–2024). Warmer colors indicate larger increases, while blue indicates places where the estimated WFH share declines.

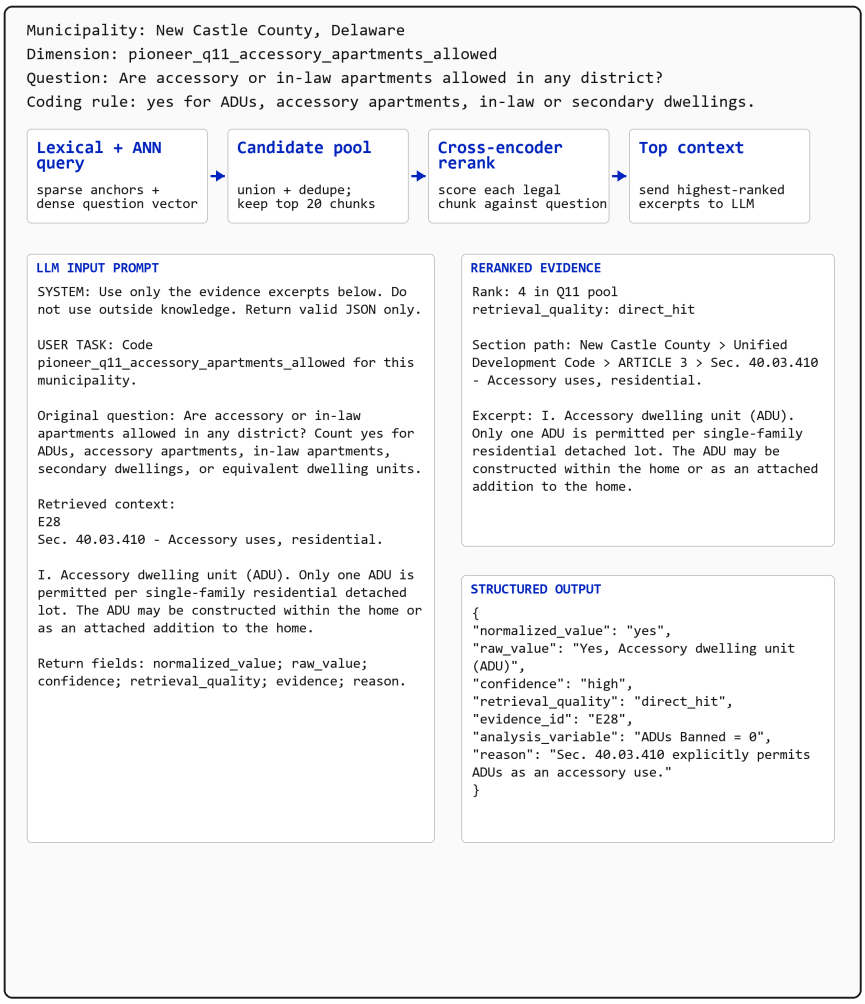

Fig. A2: Example RAG coding trace for an accessory dwelling unit (ADU) zoning dimension. Hybrid lexical and semantic retrieval first identifies candidate legal chunks, using Qwen3-Embedding-8B for dense embeddings; Qwen3-Reranker-8B then reranks the excerpts; and Qwen3.6-35B returns a structured code grounded in the retrieved source text.

Table A1. Descriptive statistics and coding outcomes. Panels A–B use ACS five-year estimates; Panels C–D summarize zoning-code extraction and embedding metadata. The main sample consists of matched Census places with at least 1,000 workers in the ACS windows ending in 2019 and 2024. PCA variance is computed descriptively on the matched embedding sample; prediction models fit PCA on training folds only.

Variable	*N*	Mean	SD	Median
Panel A. Outcome				
Δ WFH share, 2019–2024 (pp)	4,650	8.376	6.515	7.293
Panel B. Pre-pandemic baseline variables				
WFH share, 2019	4,650	0.047	0.034	0.040
Broadband share, 2019	4,650	0.823	0.092	0.833
No internet share, 2019	4,650	0.143	0.082	0.132
Bachelor's+ share, 2019	4,650	0.309	0.171	0.263
Remote-friendly industry share, 2019	4,650	0.124	0.074	0.107
Computer/engineering/science occupation share, 2019	4,650	0.051	0.034	0.044

Panel C. Zoning-code numeric outcomes

Dimension	*n*	**Code distribution (%)**
Direct margin: home–work legality		
Home-occupation permission	3,345	0=prohib. 0.2; 1=CUP/SP 14.0; 2=admin. 15.5; 3=by right 70.4
Approval friction	3,252	0=disc. 23.2; 1=admin. 20.8; 2=by right 55.9
Nonresident-employee cap	2,574	0=none 57.8; 1=any 42.2
Client/customer visits	2,953	0=prohib. 10.6; 1=limited 83.1; 2=allowed 6.2
Floor-area cap	2,231	0=cap 96.7; 1=no cap 3.3
Sales, storage, shipping	2,986	0=prohib. 47.7; 1=limited 51.4; 2=permitted 0.9
Externality standards	3,361	0=specific standards 98.9; 1=no specific standards 1.1
Indirect margin: neighborhood adaptation		
ADU permission	4,526	0=not allowed 68.1; 1=allowed 31.9
Mixed-use/live-work	4,650	0=not allowed 54.0; 1=allowed 46.0
Apartments above commercial	4,650	0=no evidence 81.3; 1=positive evidence 18.7
Off-street parking minimums	3,953	0=required $>$ 99; 1=none $<$ 1
Parking flexibility	4,650	0=not allowed 50.3; 1=allowed 49.7

Panel D. Zoning-text embedding representation

Item	**Value**
Places with zoning-code embeddings	4,650
Original embedding dimension	4,096
Chunks per place, mean	233.3
Chunks per place, median	194
Chunks per place, IQR	105–308
PCA-50 cumulative variance explained	69.8%
PCA-100 cumulative variance explained	79.8%

Table A2. Extraction questions for the twelve zoning dimensions

Dimension	Original extraction question
Direct margin: home–work legality	
Home-occupation permission	Is a home occupation allowed in residential districts, and at what approval level?
Approval friction	Is home occupation allowed only through discretionary review (conditional/special permit)?
Nonresident-employee cap	Are nonresident employees allowed, and is there a cap?
Client/customer visits	Are client visits allowed, limited, or prohibited?
Floor-area cap	Is there a floor-area or percentage cap on home-occupation space?
Sales, storage, shipping	Are sales/storage/shipping for the home business allowed or restricted?
Externality standards	Is it governed mainly by externality/performance standards (noise, traffic, signs, parking)?
Indirect margin: neighborhood adaptation	
ADU permission	Are accessory or in-law apartments allowed in any district? If allowed by special permit in any district, count as allowed.
Mixed-use/live-work	Does the code allow mixed-use or live-work arrangements?
Apartments above commercial	Are dwellings above commercial or retail allowed?
Off-street parking minimums	Does the code impose off-street parking minimum requirements?
Parking flexibility	Does the code allow parking flexibility (reductions, shared parking, waivers, alternatives)?

Table A3. Zoning provisions and growth in work-from-home share. The dependent variable is the change in ACS five-year WFH share from 2019 to 2024, in percentage points. Each row reports a separate complete-case regression of ΔWFH on the listed zoning measure and ACS controls, with state fixed effects added in the second column of each block. Coefficients report percentage-point changes associated with a one-standard-deviation increase in the zoning measure. Standard errors clustered by state are in parentheses. The analytic sample size n is reported per row because each provision is estimated in a separate complete-case regression. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

		Home-work legality		Neighborhood adaptation	
Variable	n	**ACS**	**+FE**	**ACS**	**+FE**
Direct margin: home-work legality					
Home-occupation permission	3,345	0.031	0.068		
		(0.079)	(0.062)		
Approval friction	3,252	-0.007	0.092		
		(0.072)	(0.062)		
Nonresident-employee cap	2,574	0.008	0.011		
		(0.094)	(0.078)		
Client/customer visits	2,953	-0.142*	-0.080		
		(0.081)	(0.078)		
Floor-area cap	2,231	0.048	0.010		
		(0.067)	(0.074)		
Sales/storage/shipping	2,986	0.002	-0.030		
		(0.083)	(0.088)		
Externality standards	3,361	-0.079	-0.047		
		(0.053)	(0.050)		
Indirect margin: neighborhood adaptation					
ADU permission	4,526			0.444***	0.222**
				(0.086)	(0.089)
Mixed-use/live-work	4,650			0.385***	0.250***
				(0.096)	(0.090)
Apartments above commercial	4,650			0.105	0.075
				(0.068)	(0.052)
No off-street parking minimum	3,953			0.024	0.006
				(0.087)	(0.078)
Parking flexibility	4,650			0.374***	0.230***
				(0.090)	(0.088)
ACS baseline controls		Yes	Yes	Yes	Yes
State fixed effects		No	Yes	No	Yes

Table A4. Predictive content of zoning text for WFH growth. Entries report held-out R^2 from an 80/20 train–test split. Rows containing the ACS baseline include state fixed effects. Zoning-code embeddings are 4,096-dimensional mean-pooled document embeddings reduced to 50 or 100 principal components, with PCA fit on the training split only. HGB denotes histogram gradient boosting.

Feature set	OLS	Lasso	HGB	XGBoost
Embedding only, PCA-50	0.230	0.228	0.238	0.204
Embedding only, PCA-100	0.293	0.287	0.291	0.223
ACS baseline + state FE	0.606	0.606	0.609	0.613
+ 12 RAG dimensions	0.609	0.610	0.616	0.617
+ embedding (PCA-50)	0.630	0.628	0.628	0.623
+ embedding (PCA-100)	0.632	0.629	0.629	0.617
+ 12 RAG + embedding (PCA-100)	0.633	0.628	0.631	0.628