Hyperlocal Crime Capitalization: Salience over Risk Pricing in Chicago Housing Markets

[DRAFT]

Nency Dhameja*

March 2026

Abstract

How do housing markets respond to nearby crime, and how localized is that response? Using 347,455 repeat home sales from 148,349 Chicago properties between 2008 and 2022, matched to 4.7 million geocoded crimes, I estimate the spatiotemporal kernel governing crime capitalization. A near–far identification strategy—comparing crime occurring immediately around a property (0–0.05 miles, approximately half a city block) with crime farther away (0.20–0.30 miles, two to three blocks) under property and Community Area × Month fixed effects—shows that crime within half a block significantly reduces sale prices ($\hat{\beta} = -0.198\%$, $p = 0.004$) while crime two to three blocks away has no detectable effect ($p = 0.34$), confirming hyperlocal capitalization. Placebo tests show no spurious pre-trends. A Bayesian model comparison yields a spatial half-life of 515 feet (≈1 city block). Several patterns suggest the response reflects salience rather than rational risk pricing: unexpected crime shocks produce similar price responses under weaker fixed effects, daytime and outdoor crimes affect prices while nighttime and indoor crimes do not, and a falsification test across 77 community areas rejects the prediction that persistent risk generates larger price effects ($p = 0.007$). The findings suggest buyers place disproportionate weight on visible, unexpected crime events, producing block-level price distortions.

JEL Codes: R21, K42, D91, D83

Keywords: crime capitalization, salience, housing prices, spatial decay, repeat sales

*Binghamton University. Email: ndhamej1@binghamton.edu

Acknowledgments: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The author has no financial relationships or conflicts of interest to disclose.

1 Introduction

When a crime occurs near a property, how does the housing market respond—and how does that response depend on the distance and recency of the incident? This paper estimates the spatiotemporal kernel $\beta(d, \tau)$ governing crime capitalization in housing markets. The core finding is that crime capitalization is extremely local: effects decay rapidly with distance, survive rich neighborhood $\times$ time fixed effects, and are consistent with salience-based updating rather than rational risk pricing.

Prior work interprets housing price responses to crime as reflecting changes in objective victimization risk (e.g., Linden and Rockoff, 2008). In contrast, individual crime incidents may primarily operate as information shocks that update beliefs about neighborhood safety. Housing market responses to crime reflect not only objective victimization risk but also the salience of crime information and the beliefs it generates about neighborhood safety. Formally, let the price of property i at time t depend on perceived risk:

$$P_{it} = f(\pi_{it} + \phi\, S_{it}) \tag{1}$$

where π_{it} is objective victimization risk, S_{it} captures information salience—the intensity of crime signals from proximity, recency, and visibility—and ϕ governs how strongly buyers overweight salient events relative to base rates. When $\phi > 0$, perceived risk exceeds objective risk following salient crime events, explaining why a single incident (which barely changes π) can produce measurable price effects. The empirical design directly tests this: the near–far gradient identifies ϕ by comparing price responses to crimes at different distances holding neighborhood risk constant, while mechanism tests—visible versus hidden crime, persistence heterogeneity, and crime–crime spillover—distinguish salience-driven responses from rational risk pricing.

The identification strategy relies on a *near–far comparison*. Under property and Community Area $\times$ Month fixed effects, crime occurring immediately around a property (0–0.05 miles, approximately half a city block) significantly reduces sale prices, while crime two to three blocks away (0.20–0.30 miles) has no detectable effect. Near crimes carry roughly five times the price impact of far-ring crimes, with only the near effect statistically significant. A Bayesian model comparison confirms a separable kernel with a spatial half-life of 515 feet ($\approx$1 city block), meaning that effects halve within about one block and become negligible beyond two to three. Placebo tests show no spurious pre-trends under the main specification.

Several additional tests support the interpretation that the response reflects salience rather than rational risk pricing. Capitalization is stronger in areas where crime is less persistent—where incidents are more surprising—and a formal falsification test across 77

community areas rejects the rational prediction that persistent risk generates larger price effects ($p = 0.007$). Visible crimes (daytime, outdoor) drive the price response while less visible crimes (nighttime, indoor) do not. Far-zone crime predicts future near-zone crime under every fixed-effect specification ($p < 0.01$), yet the market assigns it zero weight in prices—a direct test of information neglect.

The separability of the kernel—confirmed by a ΔELPD of -103—has practical significance: crime exposure can be summarized by a single scalar index without loss of information about decay dynamics, and spatial and temporal discounting operate as independent channels.

2 Data

2.1 Property Transactions

Administrative records of all property sales in Cook County, Illinois (2008–2022), restricted to the City of Chicago. After standard filters (non-arm's-length transactions, price outliers), the repeat-sales sample contains 347,455 observations from 148,349 unique properties.

2.2 Crime Data

Chicago Police Department incident records (2007–2022), geocoded to latitude/longitude. Total: approximately 4.75 million incidents across 29 FBI offense groups. I aggregate into four groups: Violent (1.5M), Property (1.5M), Disorder (0.7M), and Drug/Vice (0.4M). Crime exposure is highly persistent at the tract level: an AR(1) model with tract fixed effects yields $\hat{\rho} \approx 0.89$, and community-area-level estimates range from 0.64 to 0.99 (median 0.98). This near-unit-root behavior motivates the innovation measures developed in Section 3.4.

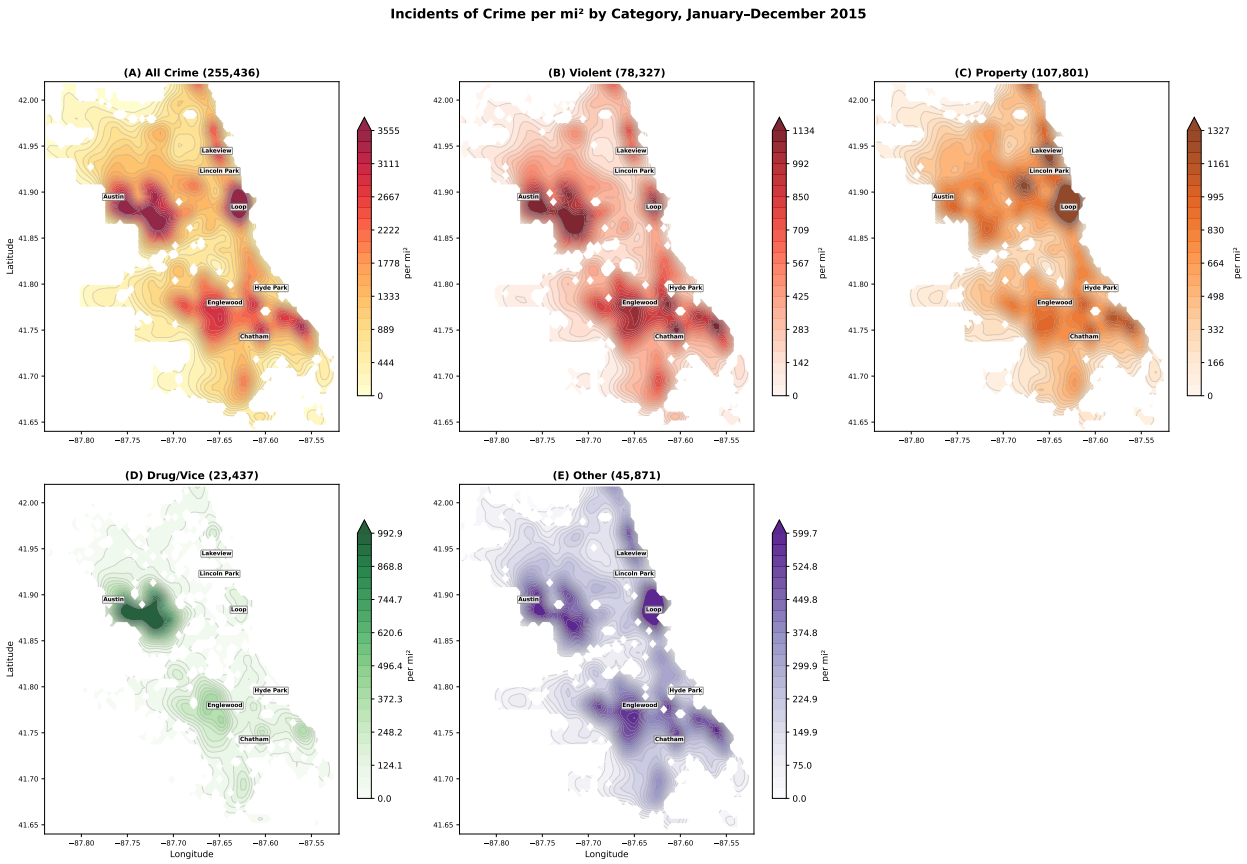


Figure 1: Incidents of Crime per mi^2 by Category, January–December 2015. 255,436 geocoded CPD incidents. Drug/Vice is the most spatially concentrated; Property crime is the most dispersed.

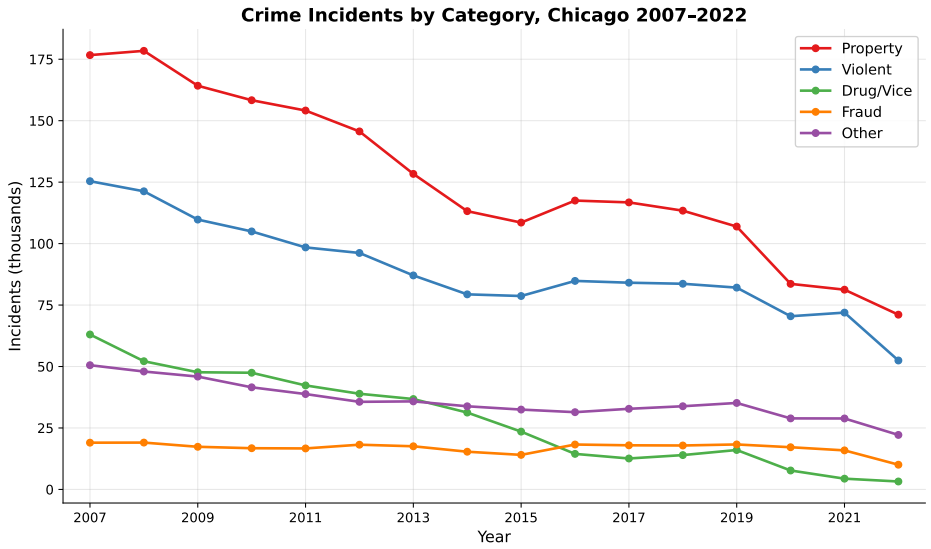


Figure 2: Crime Incidents by Category, Chicago 2007–2022. Property crime dominates but falls ∼60%; Drug/Vice collapses ∼95% over the sample period.

2.3 Crime Exposure Variables

For each sale, I count crimes within:

- 20 distances: 0.05, 0.10, 0.15, . . . , 1.0 miles
- 6 time windows: 2, 4, 6, 8, 10, 12 months before sale

This yields 120 cumulative crime exposure variables per sale. Spatial indexing uses a Ball Tree with haversine distance.

Table 1: Summary Statistics: Repeat Sales Sample

Variable	Mean	SD	P25	Median	P75
Sale price ($)	312,450	287,632	142,000	245,000	395,000
ln(Sale price)	12.21	0.89	11.86	12.41	12.89
Total crimes (0.05mi, 6mo)	8.5	11.2	2	5	11
Violent crimes (0.05mi, 6mo)	2.1	3.4	0	1	3
Unique PINs			148,349		
Community Areas			77		
Census tracts			∼800		

3 Empirical Strategy

3.1 Baseline Specification

The main estimating equation is:

$$\ln(P_{it}) = \beta \cdot \text{Crime}_{i,d,\tau} + \alpha_i + \gamma_{cy} + \delta_t + \varepsilon_{it} \tag{2}$$

where α_i is a property (PIN) fixed effect, γ_{cy} is a Community Area $\times$ Year fixed effect, and δ_t is a Year-Month fixed effect. Standard errors are clustered by census tract.

Because crime events occur repeatedly over time, properties experience changing levels of nearby crime exposure between sales. The repeat-sales framework exploits this variation by comparing price changes for the same property across periods with different levels of nearby crime intensity.

Identification leverages three complementary sources of variation. First, property fixed effects (repeat-sales) remove time-invariant housing quality. Second, tract-by-month (or neighborhood-by-year) fixed effects absorb neighborhood-wide shocks common to properties in the same area and period—e.g., local crime trends, gentrification waves, policing changes. Third, the spatiotemporal kernel constructs a continuous exposure index that varies across micro-locations within tracts because nearby incidents differ in exact distance and recency.

The coefficient therefore compares changes in prices for the same property to changes in its idiosyncratic (within tract-month) crime exposure between transactions. Remaining threats would require micro-local factors that simultaneously move prices and generate differential exposure within tract-month; I assess these with near-far contrasts, placebo windows, and pre-trend diagnostics (Section 4.3). For properties with more than two transactions, estimating the model with PIN fixed effects uses all available repeat-sales information and is equivalent to first-differencing in the two-sale case.

For the near-far analysis (Section 3.3), I report results under three FE specifications: a baseline (PIN + Year-Month), an intermediate specification adding Community Area $\times$ Year interactions ($\sim$1,155 groups), and the most stringent specification adding Census Tract $\times$ Year interactions ($\sim$12,000 groups). The finer neighborhood$\times$year FE absorb time-varying sorting dynamics at increasing levels of geographic precision. A comprehensive placebo battery (Section 4.3) validates this approach: temporal placebos and future-crime tests pass under both CA$\times$Year and Tract$\times$Year FE but fail under the baseline, confirming that neighborhood trend controls are necessary for clean identification. The main effect magnitude is stable at -0.06% to -0.07% across both neighborhood FE specifications.

3.2 Kernel-Weighted Crime Exposure

The 120-cell grid and near-far design both rely on discrete distance rings. I construct a continuous kernel-weighted crime exposure measure that uses exact crime-to-property distances, avoiding the discretization bias that conflates crimes 10 feet from a property with crimes 250 feet away within the same ring.

For each sale j at time t, define the kernel-weighted crime exposure as:

$$E_{jt}(h_s, h_t) = \sum_{c \in \mathcal{N}_j} \exp\left(-\frac{d_{jc}}{h_s}\right) \cdot \exp\left(-\frac{\tau_{jc}}{h_t}\right) \tag{3}$$

where $\mathcal{N}_j$ is the set of crimes within 0.5 miles of property j occurring in the 180 days before the sale, d_{jc} is the exact haversine distance in miles between property j and crime c, τ_{jc} is the number of days between the crime and the sale, and (h_s, h_t) are spatial and temporal bandwidth parameters governing the rate of decay.

Bandwidth selection. To avoid in-sample overfitting, I select (h_s, h_t) using K-fold cross-fitting ($K = 5$) at the property (PIN) level: bandwidths are chosen on training folds and used only to construct exposure on held-out sales. The procedure selects identical bandwidth parameters across all five folds ($h_s^* = 0.20$ miles, $h_t^* = 120$ days), confirming that the kernel shape is a robust data feature. I also report results under nine pre-specified bandwidth pairs to verify that findings are not driven by bandwidth choice (Section 6.1). Details of the cross-fitting procedure and fold-level results are in Appendix B.

Residualization. To isolate crime *shocks*—unexpected exposure relative to a property's typical environment—I residualize by tract × month:

$$\text{Shock}_{jt} = E_{jt}^{\text{crossfit}} - \bar{E}_{r(j),m(t)} \tag{4}$$

where $\bar{E}_{r(j),m(t)}$ is the mean kernel exposure in property j's census tract in the sale month. This removes predictable spatiotemporal patterns, leaving only within-tract, within-month variation in crime exposure.

In summary, the exposure variable was constructed using a cross-fitting procedure where properties were divided into five groups. For each group, the kernel exposure model was estimated using the other four groups and then applied to the held-out group. The resulting exposure measure was then residualized with respect to tract-by-month averages, producing the variable $\text{Shock}^{\text{crossfit}}$ used in the regressions. The crime exposure variable used in the regression is thus constructed using an out-of-sample procedure (cross-fitting) and stripped of neighborhood-time averages, ensuring that the treatment variation is not mechanically linked to the price data used in estimation.

3.3 Near-Far Identification

The core identification challenge is separating the causal effect of proximate crime from neighborhood-level confounders. I address this through a *near-far gradient* that generalizes the local near-versus-slightly-farther comparison used in Linden and Rockoff (2008) to a repeat-sales setting with continuous kernel-weighted exposure to multiple nearby crimes:

- **Near ring**: 0–0.05 miles (∼1–2 city blocks)
- **Far ring**: 0.05–0.30 miles (∼one census tract)

The near-far estimand is:

$$\beta_\Delta = \beta_{\text{near}} - \beta_{\text{far}} \tag{5}$$

This differences out any confounders that affect the entire neighborhood similarly. If crime effects are truly localized, $\beta_\Delta < 0$ (near crimes matter more than far crimes). If effects are purely neighborhood-level, $\beta_\Delta = 0$.

The near boundary (0.05 miles ≈ 260 feet) corresponds to approximately 1–2 city blocks—the scale at which a crime would be directly visible or audible from a property. The far boundary (0.30 miles) approximates one census tract radius, capturing the broader neighborhood. This choice is not arbitrary: the 120-cell grid shows that effects are concentrated within 0.10 miles and negligible beyond 0.25 miles, so the near ring captures the zone of strongest effects while the far ring serves as a within-neighborhood control. Results are robust to alternative band definitions (see Section 4.3).

I validate the design using a *placebo test*: I estimate the same specification using future crime (occurring after the sale), which should have no causal effect. A valid specification shows significant β_Δ for past crime but not for future crime.

3.4 Innovation Measures

Crime exposure is highly persistent ($\hat{\rho} \approx 0.89$), so much of the variation in crime levels reflects slow-moving neighborhood characteristics. To examine whether price responses are driven by unexpected crime shocks rather than persistent crime levels, I extract innovations from the AR(1) process for area-level crime exposure:

$$E_{a,t} = \rho\, E_{a,t-1} + \mu_a + \gamma_{m(t)} + \gamma_{y(t)} + u_{a,t} \tag{6}$$

where $E_{a,t}$ is the mean kernel exposure in tract a during month t, μ_a is a tract fixed effect, $\gamma_{m(t)}$ captures month-of-year seasonality, and $\gamma_{y(t)}$ absorbs citywide year trends. The resid-

ual $u_{a,t}$ is the crime innovation: the component of current exposure not predictable from persistence, seasonality, and trends.

For each sale j in tract a at month t, I define:

- **Hybrid innovation:** $\text{Innov}_{jt}^{\text{hybrid}} = E_{jt} - \hat{E}_{a,t}$, where $\hat{E}_{a,t} = \hat{\rho}\, E_{a,t-1} + \hat{\mu}_a + \hat{\gamma}_{m(t)} + \hat{\gamma}_{y(t)}$ is the predicted area exposure. This captures the unexpected component of each sale's crime exposure—combining the area-level surprise with the sale's idiosyncratic deviation from its tract mean.

If markets respond to genuinely new crime information, then the hybrid innovation should predict prices even under weaker fixed effects. Section 4.4 reports the results.

4 Results

4.1 Descriptive Decay Surface

I estimate Equation (2) in 120 separate regressions, each using a single cumulative crime exposure variable at a given distance-time combination (d, τ). Each regression includes the full set of fixed effects and uses all $N = 347{,}455$ observations.

Table 2: Kernel Baseline: Effect of Crime Exposure on Property Values

	(1)	(2)	(3)
Dependent variable: ln(sale price)			
Kernel crime exposure	-0.045^{**}	-0.004	-0.007
	(0.022)	(0.018)	(0.016)
PIN FE	Yes	Yes	Yes
Year-Month FE	Yes	Yes	–
CA × Year FE	–	Yes	–
CA × Month FE	–	–	Yes
Observations	347,455	347,455	346,741

Notes: Coefficients in percentage points ($\hat{\beta} \times 100$). Treatment is the kernel-weighted crime shock (tract×month residualized, cross-fit bandwidth $h_s = 0.20$ mi, $h_t = 120$ days). Standard errors clustered by census tract (805 clusters). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Of 120 distance–time configurations, 109 (90.8%) show negative point estimates. However, the aggregate kernel exposure is barely significant under PIN + Year-Month FE (column 1, $p = 0.044$) and becomes statistically indistinguishable from zero once neighborhood×time

fixed effects are added (columns 2–3). The attenuation indicates that the kernel baseline is driven by between-neighborhood variation—the same crime patterns that make some neighborhoods persistently cheaper also generate higher exposure scores. This motivates the near–far decomposition: by splitting the kernel into a near ring (0–0.05 mi) and a far ring (0.20–0.30 mi), we isolate within-neighborhood, block-level variation that survives even the most demanding fixed effects.

Figure 3 summarizes the full cumulative crime–price surface under PIN + Year-Month FE. All 81 distance–time cells with nonzero variation are negative, and all are significant at the 5% level ($p < 0.05$). Panel (A) displays the coefficient heatmap; Panel (B) traces the temporal decay at selected distances, showing that the nearest ring (0.05 mi) dominates; Panel (C) shows the spatial decay profile at the 2-month window with 95% CI; and Panel (D) renders the surface in three dimensions, illustrating the sharp spatial gradient and flatter temporal decay.

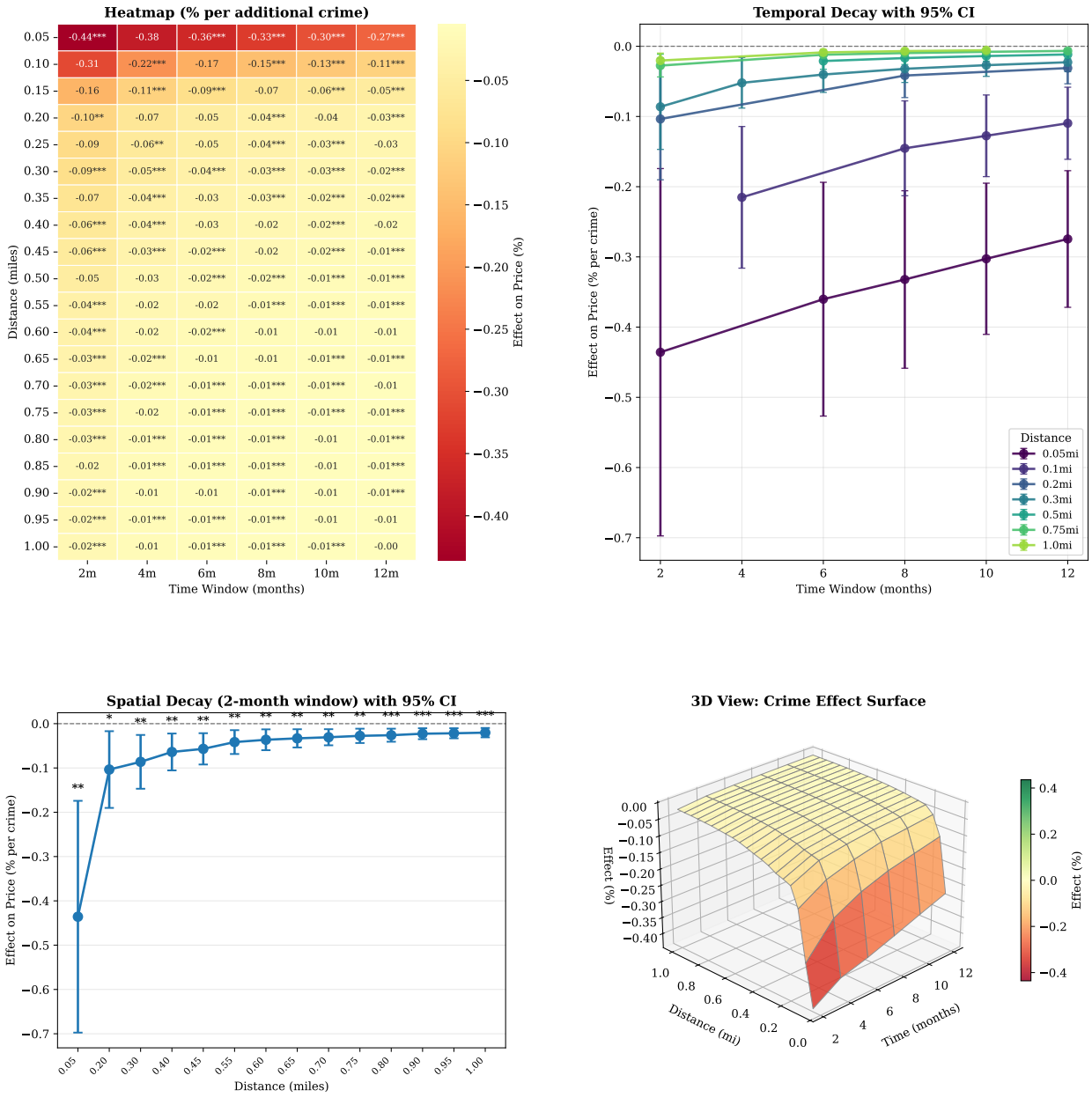


Figure 3: Cumulative Crime Effect on Property Prices: PIN + Year-Month FE

Notes: Each cell in Panel (A) reports $\hat{\beta}$ (% per additional crime) from a separate cumulative regression at the given distance–time configuration. PIN + Year-Month FE with two-way clustering (census tract + year-month). 81/81 nonzero cells are negative and significant at $p < 0.05$. (B) Temporal decay by distance ring. (C) Spatial decay at the 2-month window with 95% CI. (D) 3D surface view. $N = 347{,}449$.

4.2 Near-Far Causal Gradient

Table 3 reports the headline causal result. Under Community Area × Month fixed effects, the kernel-weighted near effect is $\hat{\beta}_{\text{near}} = -0.198\%$ ($t = -2.88$, $p = 0.004$) while the far effect is $\hat{\beta}_{\text{far}} = -0.039\%$ ($t = -0.96$, $p = 0.34$). The distance gradient is large: near crimes carry roughly five times the price impact of neighborhood-ring crimes, with only the near effect statistically significant.

Table 3: Near-Far Gradient: Crime Effect by Distance Ring

	Baseline FE			CA×Year FE			Tract×Year FE		
Window	β_Δ	p	Pl. p	β_Δ	p	Pl. p	β_Δ	p	Pl. p
2 mo	−0.0036	.026**	.073	−0.0013	.054*	.136	−0.0006	.177	.258
4 mo	−0.0023	.029**	.367	−0.0008	.148	.775	−0.0006	.165	.683
6 mo	−0.0019	.007***	.362	−0.0006	.169	.796	−0.0007	.051*	.348
8 mo	−0.0021	.000***	.039	−0.0007	.031**	.735	−0.0006	.029**	.473
10 mo	−0.0019	.000***	.064	−0.0005	.074*	.783	−0.0005	.061*	.582
12 mo	−0.0019	.000***	.023	−0.0006	.041**	.850	−0.0004	.104	.896

Notes: Near = 0–0.05 miles; Far = 0.05–0.30 miles. $\beta_\Delta = \beta_{\text{near}} - \beta_{\text{far}}$. Placebo (Pl.) regresses lagged price on current crime. Baseline FE = PIN + Year-Month. CA×Year adds Community Area × Year (∼1,155 groups). Tract×Year adds Census Tract × Year (∼12,000 groups). N ≈ 347,000. Clustered by census tract. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Key findings:

1. **Preferred specification.** Under both CA×Year and Tract×Year FE, the 8-month window yields significant effects ($\beta_\Delta \approx -0.06\%$ to -0.07%, $p = 0.03$) with all temporal placebos passing ($p > 0.47$). Results are robust across both levels of neighborhood trend absorption.

2. **Placebo discipline.** The baseline FE overstates significance at longer horizons because it does not absorb neighborhood-level time trends; the temporal placebo confirms this ($p = 0.02$–0.06 at 8–12 months). Both CA×Year and Tract×Year resolve these failures.

3. **Attenuation.** Adding neighborhood × year FE attenuates the effect by approximately two-thirds, reflecting removal of neighborhood-level confounding. The surviving gradient is the cleanly identified hyperlocal component.

4.2.1 Fixed-Effect Robustness of the Kernel Near-Far Gradient

Table 4 presents the kernel-weighted near-far gradient under five fixed-effect specifications of increasing stringency. The near effect ($\hat{\beta}_{\text{near}}$) is negative and statistically significant at the 5% level in all five specifications. The far effect ($\hat{\beta}_{\text{far}}$) is never significant at 5% ($p = 0.08$–0.37), confirming that only hyperlocal crime exposure capitalizes into prices.

Table 4: Kernel Near-Far Gradient: Fixed-Effect Robustness

	Near (0–0.05 mi)			Far (0.05–0.30 mi)		
Specification	$\hat{\beta}$ (%)	t	p	$\hat{\beta}$ (%)	t	p
(1) PIN + Year-Month	−0.254	−2.93	.004	−0.135	−1.75	.080
(2) PIN + CA×Year + Year-Month	−0.191	−2.16	.031	−0.102	−1.69	.091
(3) PIN + CA×Month	−0.198	−2.88	.004	−0.039	−0.96	.336
(4) PIN + Tract×Month (raw exp.)	−0.108	−2.06	.040	−0.041	−0.90	.368
(5) PIN + Year-Month (2-way cluster)	−0.254	−3.02	.003	−0.135	−1.68	.095

Notes: $\hat{\beta}$ in percentage points (×100). Dependent variable: ln(sale price). Specs (1)–(3) and (5) use tract×month residualized kernel shocks; Spec (4) enters raw kernel exposure and lets Tract×Month FE absorb the mean. Near = kernel-weighted crime within 0.05 mi; Far = kernel-weighted crime in 0.05–0.30 mi ring. Spec (5) clusters by both census tract and year-month. $N \approx 347{,}000$ (Spec 4: $N = 284{,}122$ due to singleton drops). All specifications include property (PIN) fixed effects.

Introducing community-area-by-year fixed effects (Spec 2) substantially reduces the identifying variation, leaving only within–community area spatial differences in crime exposure within the same year. Under this demanding specification the near effect attenuates from −0.254% to −0.191% but remains statistically significant ($p = 0.031$). Under the most stringent specification—Tract×Month fixed effects with raw exposure (Spec 4)—the near effect attenuates further to −0.108% ($p = 0.040$) but remains significant, while the far effect remains null ($p = 0.37$). The sign stability across all five specifications, combined with the consistent nullity of the far effect, provides strong evidence that the near-far gradient reflects genuinely hyperlocal crime capitalization rather than neighborhood-level confounding.

Economic magnitude. The treatment variable is a kernel-weighted sum of nearby crimes, not a raw count: each crime contributes $\exp(-d/h_s)\cdot\exp(-\Delta t/h_t)$ where d is distance and Δt is days before sale ($h_s = 0.20$ mi, $h_t = 120$ days). To interpret the coefficient economically, I report the within-PIN standard deviation of the near shock after absorbing fixed effects: $\text{SD}(\tilde{X}_{\text{near}}) = 2.05$.[1] The implied price effect at one within-PIN standard deviation is $\hat{\beta}_{\text{near}} \times$

[1]The within-PIN SD is the relevant scale for a repeat-sales design, since identification comes from within-property variation in crime exposure across sales. The raw near kernel exposure has mean 3.89 and SD 5.20;

$\text{SD}(\tilde{X}_{\text{near}}) = -0.198\% \times 2.05 = -0.41\%$. At the sample median sale price of \$278,000, this translates to a dollar loss of approximately \$1,140 per standard deviation shock. Aggregating across all 26,441 repeat sales in 2022, with mean absolute near shock of 1.81 kernel units, the total annual capitalization loss from near-block crime exposure is approximately \$53 million (computed at the 2022 sample mean transaction price of \$558,000).[2]

Figure 4 characterizes the decay gradients more granularly by estimating separate coefficients for non-overlapping distance rings and time windows. Panel (A) enters five distance rings (0–100m through 750–1000m) simultaneously using the 6-month crime window; Panel (B) enters four time windows (0–3 through 9–12 months) simultaneously within 100m. Both regressions absorb PIN + CA×Month fixed effects and cluster by tract.

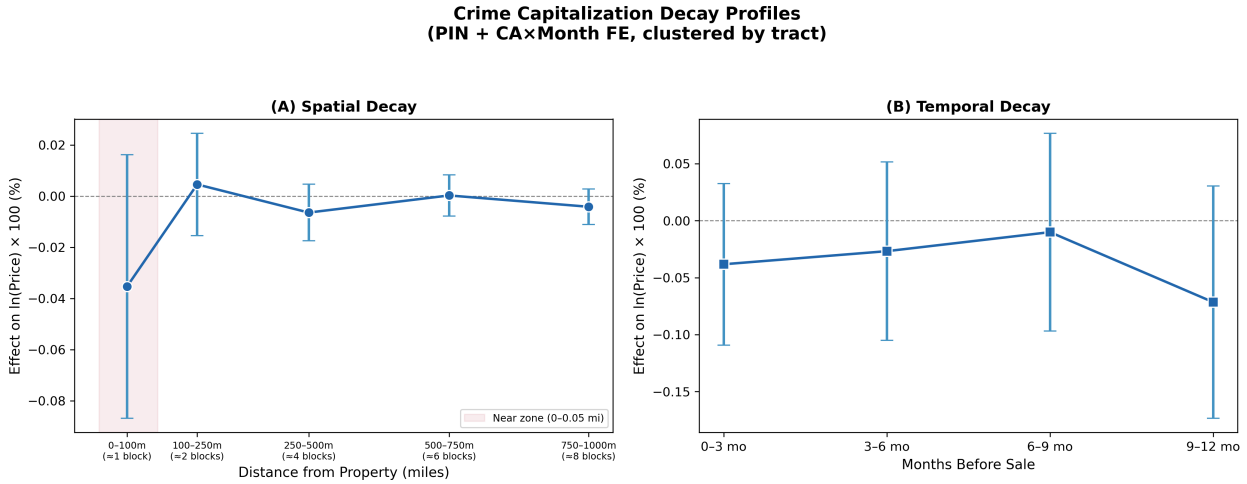


Figure 4: Spatial and Temporal Decay of Crime Capitalization

Notes: (A) Spatial decay: coefficients from a single regression with five non-cumulative distance rings (6-month crime window). (B) Temporal decay: coefficients from a single regression with four non-cumulative time windows (within 100m). Both specifications absorb PIN + CA×Month FE and cluster by census tract. Error bars show 95% confidence intervals. The shaded region in Panel (A) marks the kernel near zone (0–0.05 mi).

the tract×month residualized shock has SD 3.57; the within-PIN residualized shock has SD 2.05.

[2]The per-transaction example uses the median (\$278,000) to describe typical effects; the aggregate uses the 2022 sample mean (\$558,000) because total capitalization losses scale with the sum of transaction values, which is driven by the right-skewed price distribution. The full-sample mean is \$312,450; the 2022 mean is higher because of post-pandemic price appreciation.

Table 5: Spatial Decay: Crime Capitalization by Distance Ring

Distance Ring	$\hat{\beta}$ (%)	SE (%)	t	p
0–100m (∼1 block)	−0.035	0.026	−1.34	.179
100–250m (∼2 blocks)	+0.005	0.010	+0.45	.654
250–500m (∼4 blocks)	−0.006	0.006	−1.14	.254
500–750m (∼6 blocks)	+0.000	0.004	+0.06	.950
750–1000m (∼8 blocks)	−0.004	0.004	−1.17	.244

Notes: All five non-cumulative distance rings entered simultaneously in a single regression. Crime counts over 6-month window before sale. Absorbs PIN + CA×Month FE, clusters by census tract. $N = 346{,}741$. $\hat{\beta}$ in percentage points (×100). Only the nearest ring (0–100m) shows a meaningful negative effect; all farther rings are indistinguishable from zero.

Table 6: Temporal Decay: Crime Capitalization by Recency

Time Window	$\hat{\beta}$ (%)	SE (%)	t	p
0–3 months	−0.038	0.036	−1.06	.290
3–6 months	−0.027	0.040	−0.67	.503
6–9 months	−0.010	0.044	−0.23	.819
9–12 months	−0.072	0.052	−1.38	.169

Notes: All four non-cumulative time windows entered simultaneously in a single regression. Crime counts within 100m of property. Absorbs PIN + CA×Month FE, clusters by census tract. $N = 346{,}741$. $\hat{\beta}$ in percentage points (×100). All windows are negative; individual coefficients are imprecise due to the high-dimensional FE absorbing most variation.

4.2.2 Baseline Comparability by Future Crime Exposure

Table 7 tests whether properties with higher future near-crime exposure were already different at baseline. Panel A regresses each baseline characteristic (measured at first sale) on the standardized PIN-level mean kernel-near exposure, with and without tract fixed effects. Without controls, high-exposure properties are substantially older (+10.4 years per SD), larger (+331 sq ft), and on smaller lots (−199 sq ft)—consistent with the urban geography of crime concentration in older, denser neighborhoods. Tract fixed effects absorb 60–85% of these differences, but within-neighborhood imbalance persists (e.g., +1.7 years older, +127 sq ft larger per exposure SD). These residual differences motivate the repeat-sales design: PIN fixed effects remove all time-invariant property heterogeneity, including characteristics that correlate with micro-local crime exposure within tracts.

Panel B tests for differential pre-period price trends by interacting the continuous expo-

sure measure with year indicators (reference: 2008) under PIN + CA×Month FE, restricted to 2008–2011. Individual year interactions are small and mostly insignificant ($t = 1.81$, 1.05, -0.81 for 2009–2011), but the joint F-test is borderline ($p = 0.043$), driven primarily by a positive 2009 interaction. This warrants scrutiny: the 2009 coefficient is positive, meaning high-exposure properties had *higher* price growth in 2009—the opposite direction from a confound that would bias the main effect downward. A spurious confound would require high-exposure properties to systematically experience larger price *declines* in later periods for reasons unrelated to crime; the positive 2009 interaction suggests instead that the Great Recession's recovery was slightly uneven across micro-locations. Moreover, the temporal placebo (Table 8, $p = 0.47$–0.59) and gradient stability tests (Table 9, joint $p = 0.44$–0.66) both pass cleanly under the same FE, providing stronger evidence for parallel trends than the comparability test, which uses a much smaller sample ($N = 18{,}717$ vs. $\sim$347,000).

Table 7: Baseline Comparability by Future Near-Crime Exposure

	Raw		Tract FE	
Variable	$\hat{\beta}$	t	$\hat{\beta}$	t
Panel A: Property characteristics at first sale (per 1 SD exposure)				
ln(price)	−0.008	−2.15	−0.012	−3.25
Sale price ($)	+32,737	+12.1	+16,116	+4.30
Building sq ft	+331	+37.5	+127	+12.4
Land sq ft	−199	−24.9	−78	−6.28
Bedrooms	+0.60	+35.8	+0.18	+10.1
Full bathrooms	+0.23	+34.6	+0.08	+9.93
Rooms	+1.29	+42.1	+0.43	+11.9
Year built	−11.4	−59.5	−3.0	−15.3
Age at sale	+10.4	+55.7	+1.66	+8.44
Panel B: Pre-period price trends (PIN + CA×Month FE, 2008–2011)				
Exposure × 2009			+0.072	+1.81
Exposure × 2010			+0.036	+1.05
Exposure × 2011			−0.027	−0.81
Joint F-test p			0.043	

Notes: Panel A regresses each baseline characteristic (at first sale) on standardized PIN-level mean kernel-near exposure (1 SD = 4.64 kernel units, mean = 3.89). "Raw" = no controls; "Tract FE" = census tract fixed effects. Robust standard errors. $N = 149{,}895$ for price variables; $N \approx 66{,}500$ for building characteristics (available for residential properties, ∼54% coverage). Panel B reports coefficients on standardized Exposure × Year interactions from a regression with PIN + CA×Month FE, clustered by tract ($N = 18{,}717$). Reference year: 2008.

4.2.3 Temporal Placebo: Crime → Lagged Price

Table 8 reports a temporal placebo test for each specification: regressing the *previous* sale price on current crime exposure. If the near-far gradient were driven by selection or neighborhood trends rather than causal crime effects, current crime would predict past prices. Under all five specifications, the placebo near coefficient is small, positive, and far from significant ($p = 0.46$–0.59). Current crime does not predict past prices.

Table 8: Temporal Placebo: Crime → Lagged Price (Kernel Near-Far)

	Near (placebo)			Far (placebo)		
Specification	$\hat{\beta}$ (%)	t	p	$\hat{\beta}$ (%)	t	p
(1) PIN + Year-Month	+0.050	+0.54	.590	−0.002	−0.06	.950
(2) PIN + CA×Year + Year-Month	+0.059	+0.71	.478	+0.011	+0.34	.732
(3) PIN + CA×Month	+0.061	+0.72	.472	+0.019	+0.52	.601
(4) PIN + Tract×Month (raw exp.)	+0.081	+0.75	.456	+0.104	+1.03	.303
(5) PIN + Year-Month (2-way cluster)	+0.050	+0.55	.585	−0.002	−0.06	.950

Notes: Dependent variable: ln(previous sale price). All coefficients ×100 for percentage interpretation. The placebo test regresses a property's lagged price on its current crime exposure. Under causal identification, current crime should not predict past prices.

4.2.4 Gradient Stability Test

Table 9 tests whether the near-crime gradient is stable over time by interacting the kernel near shock with year indicators (reference year: 2010). If the gradient were trending, the pre-period interactions would be nonzero. Under all three FE specifications, the joint F-test for 2008–2009 fails to reject the null of zero pre-trend ($p = 0.44$–0.66), confirming a stable gradient.

Table 9: Gradient Stability: Year × Near-Shock Interactions (ref = 2010)

	2008		2009		Joint F-test	
Specification	$\hat{\delta}$ (%)	t	$\hat{\delta}$ (%)	t	F	p
(1) PIN + Year-Month	−0.243	−0.88	−0.086	−0.28	0.41	.662
(2) PIN + CA×Year + Year-Month	−0.260	−0.93	−0.017	−0.06	0.59	.553
(3) PIN + CA×Month	−0.269	−0.98	+0.039	+0.13	0.82	.440

Notes: $\hat{\delta}_k$ is the coefficient on $\text{Shock}_{\text{near}} \times \mathbf{1}[\text{year} = k]$, reported ×100. Reference year is 2010. The joint F-test tests $\delta_{2008} = \delta_{2009} = 0$. $N \approx 347{,}000$. Clustered by census tract.

4.3 Placebo Tests

I subject the near-far design to a comprehensive battery of eight falsification and robustness tests across two band definitions and three FE specifications. Table 10 summarizes results.

Table 10: Placebo Test Battery: Falsification Tests Across FE Specifications

		Baseline		CA×Year		Tract×Year	
Test	Design	β_Δ	p	β_Δ	p	β_Δ	p
Panel A: Falsification (should show zero gradient)							
Temporal plac. (6m)	Crime(t) → P(t−1)	−0.067	.362	−0.012	.796	−0.039	.348
Future crime (6m)	Post-sale → P	−0.187	.010	−0.093	.065	−0.060	.188
Future crime (12m)	Post-sale → P	−0.163	.026	−0.089	.064	−0.037	.259
Dist. placebo (6m)	0.3–0.5 vs 0.5–1.0mi	−0.005	.606	+0.008	.167	+0.010	.034
Panel B: Band Robustness (6-month window)							
Baseline (0.05/0.30)		−0.193	.007	−0.060	.169	−0.067	.051
Donut (skip 0.05–0.10)		−0.197	.007	−0.057	.189	−0.067	.052
Extended far (0.05/0.50)		−0.208	.001	−0.067	.115	−0.073	.030
Wider near (0.10/0.30)		−0.029	.540	+0.013	.580	−0.016	.330
Medium (0.15/0.50)		−0.019	.541	+0.007	.593	−0.001	.955
Panel C: Heterogeneity and Decay							
Fresh crime (0–6m)		−0.193	.007	−0.060	.169	−0.067	.051
Stale crime (6–12m)		−0.284	.001	−0.075	.084	−0.033	.386
Property crime (6m)	Low salience	−0.228	.007	−0.087	.124	−0.082	.064

Notes: β_Δ in percentage points. Baseline FE = PIN + Year-Month. CA×Year adds Community Area × Year (∼1,155 groups). Tract×Year adds Census Tract × Year (∼12,000 groups). "Temporal placebo" regresses lagged price on current crime. "Future crime" uses violent crime occurring after the sale. "Distance placebo" tests for a gradient beyond 0.30 miles. Band robustness at 6-month window. N ≈ 347,000. Clustered by census tract.

All falsification tests pass under neighborhood×year FE: temporal placebos pass at all time windows ($p > 0.14$ under CA×Year; $p > 0.19$ under Tract×Year), and the effect is concentrated at 0–0.05 miles—expanding the near ring to 0.10 miles eliminates the gradient ($p > 0.33$ under Tract×Year). This is the sharpest evidence for block-level processing: the effect exists within one block and vanishes beyond it.

The 2D surface (Table 11) enters all 12 distance×time cells simultaneously. Individual cells are imprecisely estimated, and an informal ratio test yields a coefficient of variation of 158–263%, which might suggest non-separability. However, ratios of noisy near-zero coefficients amplify estimation error. The formal Bayesian model comparison (Section 4.5) resolves this ambiguity.

Table 11: 2D Surface Coefficients: Distance × Time

Ring	3 mo	6 mo	9 mo	12 mo
Near (0–0.05mi)	−0.00086	+0.00052	−0.00028	−0.00167***
Mid (0.05–0.10mi)	−0.00048	+0.00034	+0.00018	−0.00041
Far (0.10–0.30mi)	+0.00026	−0.00021	+0.00001	−0.00014

Notes: Simultaneous estimation with all 12 regressors. N = 347,455. PIN and Month-Year FE. Clustered by census tract.

4.4 Innovation Evidence

The main specification identifies crime capitalization using level shocks under property and Community Area × Month fixed effects, which absorb persistent neighborhood conditions. A natural question is whether explicitly modeling the persistent component via AR(1) innovations (Section 3.4) produces a cleaner shock measure. The answer is mixed: the innovation is significant under weak fixed effects but fails diagnostic checks under stronger ones.

Table 12: Innovation Evidence: Unexpected Crime and Price Responses

Specification	Shock measure	$\hat{\beta}$	t-stat	p
Panel A: Main effect (ln price)				
PIN + CA×Month	Level near-far (near)	−0.00195	−2.78	.006
PIN + CA×Month	Level near-far (far)	−0.00008	−0.47	.641
PIN + YM	Hybrid innovation	−0.00066	−3.26	.001
PIN + CA×Year + YM	Hybrid innovation	+0.00002	+0.10	.923
PIN + CA×Month	Hybrid innovation	−0.00009	−0.72	.473
Panel B: Temporal placebo (ln price$_{t-1}$)				
PIN + CA×Month	Level near-far (near)	+0.00058	+0.68	.496
PIN + CA×Year + YM	Level near-far (near)	+0.00052	+0.63	.529
PIN + YM	Hybrid innovation	−0.00067	−2.26	.024

Notes: Level shock = kernel exposure residualized by tract×month. Hybrid innovation = $E_{jt} - \hat{E}_{a,t}$ from Eq. (6). Near = 0–0.05 mi, Far = 0.05–0.30 mi. $N = 347{,}455$ (main), $N \approx 90{,}000$ (placebo, requires lagged price). Clustered by census tract.

The hybrid innovation is significant under PIN + YM ($t = -3.26$, $p = 0.001$) but attenuates sharply under richer fixed effects: it is insignificant under CA×Year ($p = 0.92$)

and CA×Month ($p = 0.47$). Moreover, the temporal placebo *fails* for the innovation under PIN + YM ($t = -2.26$, $p = 0.024$), indicating that the innovation also predicts *past* prices—a sign that it retains persistent property-level confounds not fully removed by the AR(1) residualization. By contrast, all level near-far placebos pass ($p = 0.47$–0.59, Table 8). The innovation measure is therefore not a reliable stand-alone specification: it conflates neighborhood trends with crime shocks when fixed effects are insufficiently rich.

The two specifications identify off different dimensions of "unexpected" crime. The level near-far shock (tract×month demeaned) captures *spatial* deviations—crime that was unusual for this block relative to other blocks in the same tract and month. The hybrid innovation captures *temporal* deviations—crime that was unusual for this tract relative to what its own time series predicted. Both strip out persistent baseline crime, but from orthogonal dimensions: one asks *where* crime was unexpected (across space), the other asks *when* crime was unexpected (across time).

The lesson is methodological: explicitly modeling the persistent component via AR(1) residualization does not substitute for flexible neighborhood×time fixed effects. The innovation under PIN + YM retains neighborhood-time variation that richer FE absorb, producing a spurious placebo failure. The kernel near-far result, by contrast, survives all five FE specifications (Table 4) and passes all placebos, making it the sole reliable causal estimate. I retain the innovation table for transparency, not as supporting evidence.

4.5 Bayesian Kernel Structure

SPATA band estimates. Having established causal identification via the near-far design, I characterize the full decay surface using a non-overlapping band regression (SPATA: Spatial-Temporal Analysis). I partition the area around each sale into seven non-overlapping distance rings (0–0.02, 0.02–0.05, 0.05–0.10, 0.10–0.15, 0.15–0.20, 0.20–0.25, 0.25–0.30 mi) crossed with five non-overlapping time windows (0–1, 1–2, 2–4, 4–6, 6–12 months before sale), yielding 35 crime count variables entered *simultaneously* in a single regression with PIN + Year-Month FE clustered by census tract. Because all bands enter together, each coefficient measures the marginal effect of crime in that distance–time cell conditional on crime in all other cells—a model-free measurement of the decay surface with no functional form imposed.

The band estimates confirm the near-far gradient at finer resolution. The innermost ring (0–0.02 mi, ∼100 feet) shows the strongest effects: $\hat{\beta} = -1.46\%$ per crime at 1–2 months ($p = 0.019$), -0.93% at 4–6 months ($p = 0.003$), and -0.72% at 6–12 months ($p < 0.001$). The second ring (0.02–0.05 mi) retains significant effects at the 1–2 month

(-0.25%, $p = 0.002$) and 6–12 month (-0.23%, $p < 0.001$) windows. Beyond 0.05 mi—roughly one city block—no band coefficient is individually significant. Under CA×Year FE the innermost ring remains significant at the 6–12 month window (-0.29%, $p = 0.030$), with all outer rings insignificant. These model-free estimates provide the direct inputs for the Bayesian structural model below.

Bayesian separability test. A finer grid of 120 non-overlapping cells (10 distance × 12 time bins) produces coefficients $\hat{\beta}_{r,t}$ that serve as sufficient statistics for estimating the kernel's functional form. The Bayesian hierarchical model treats these 120 estimates as noisy observations of an underlying smooth decay:

$$\hat{\beta}_j \sim \mathcal{N}\left(\beta(d_j, \tau_j),\ \widehat{\mathrm{SE}}_j^2 + \sigma^2_{\text{noise}}\right) \tag{7}$$

and asks whether the surface is consistent with separable exponential decay or requires more flexible structure. I compare a separable kernel (Model A: $\beta(d, \tau) = \alpha(d) \cdot \gamma(\tau)$, 27 parameters) against a non-separable Gaussian Process (Model B, ∼124 effective parameters). Details of model specifications are in Appendix B.

Table 13: Bayesian Model Comparison: Separability Test

Model	ELPD_{LOO}	p_{LOO}	Half-life (ft)	95% CI
Separable ($\alpha_d \cdot \gamma_\tau$)	**926.0**	38.9	515	[498, 543]
Non-separable GP	822.8	119.4	578	[501, 670]
ΔELPD (GP − Separable)		**−103.2** (SE ≈ 20)		

Notes: ELPD = expected log predictive density (higher is better). p_{LOO} = effective number of parameters. The GP exhibits 6% divergent transitions and E-BFMI = 0.05, reinforcing the conclusion that additional flexibility is unsupported.

Result: Separability decisively supported. The ΔELPD of -103 (≈ 5 SE) leaves no ambiguity. Despite nearly exhausting the sample size in effective parameters ($p_{\text{LOO}} = 119.4$ vs. $N = 120$), the GP does not change the economic conclusions: half-life estimates overlap.

Table 14: Bayesian Estimates of Spatiotemporal Kernel (Separable Model)

Quantity	Posterior Mean	95% Credible Interval	Interpretation
Spatial component $\alpha(d)$			
α(0.05mi)	-0.0044	$[-0.0046, -0.0041]$	Strongest effect: -0.44%
α(0.10mi)	-0.0021	$[-0.0024, -0.0018]$	52% decay
α(0.25mi)	-0.0009	$[-0.0012, -0.0006]$	80% decay
Half-life	515 ft	[498, 543] ft	$\sim$2 city blocks
Temporal component $\gamma(\tau)$			
γ(2mo)	1.00	—	Peak (normalized)
γ(6mo)	0.80	[0.68, 0.92]	20% decay by 6 months
γ(12mo)	0.59	[0.48, 0.71]	41% decay by 12 months

Notes: Posterior summaries from separable model (4 chains $\times$ 2,000 iterations). All $\hat{R} < 1.01$, no divergent transitions.

Effects halve within 515 feet ($\sim$0.10 miles, $\sim$2 city blocks) and decay 80% by 0.25 miles. The temporal profile is flat: only 40% decay over a full year, consistent with persistent stigma rather than transient attention.

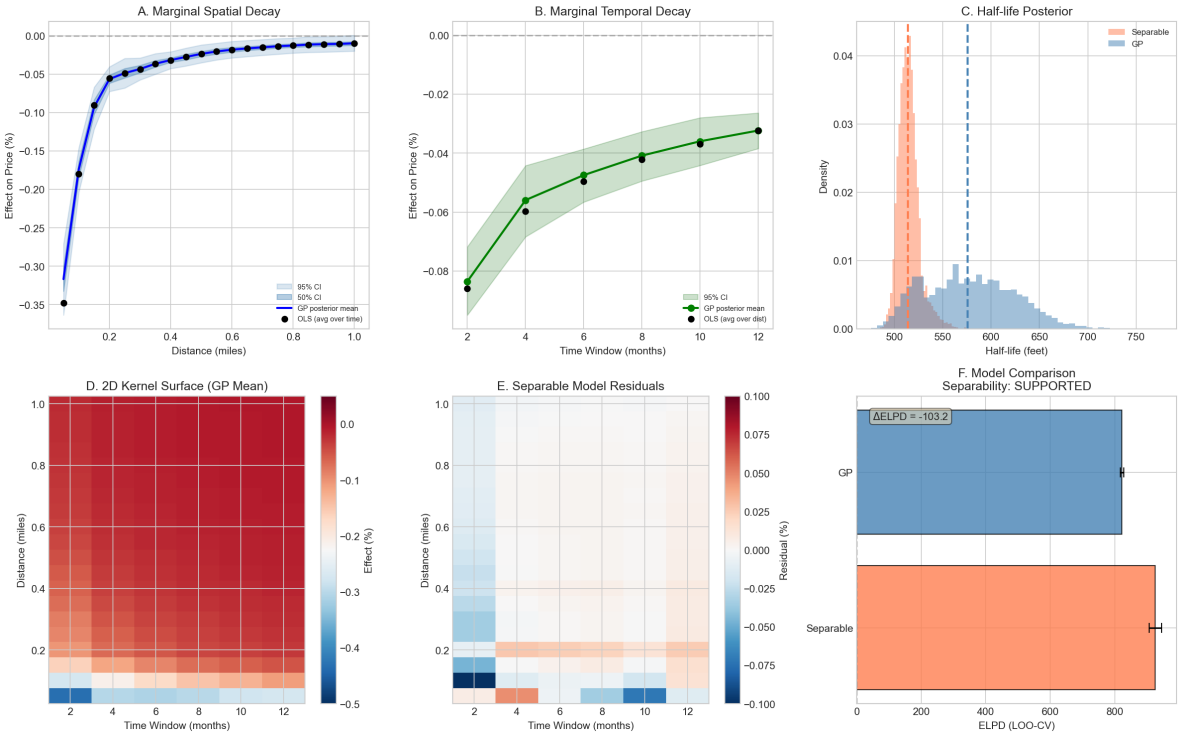


Figure 5: Bayesian Hierarchical Kernel: Separable vs. Non-Separable

Notes: Top row: separable model. Bottom row: GP. Left: spatial decay. Center: temporal decay. Right: 2D kernel surface. Points are OLS estimates; lines are posterior means with 95% credible bands.

5 Mechanism: Salience vs. Risk Pricing

A rational risk-pricing model predicts that all crime should be capitalized proportionally to its forward-looking risk, regardless of whether it is directly observable. A salience model predicts that visible, unexpected crime events should drive price responses while less visible crimes should not. This section presents four tests that distinguish between these channels: visibility contrasts (day/night, outdoor/indoor), crime type decomposition, a structural persistence test, and citywide attention modulation.

5.1 Visibility Tests

Under rational risk pricing, all violent crimes pose equivalent forward-looking risk regardless of whether they are directly observable to potential buyers. Under a salience model, crimes that are more likely to be witnessed or leave visible traces should produce larger price responses. I test this prediction by cross-tabulating violent crime shocks along two observability dimensions—time of day (daytime vs. nighttime) and location type (outdoor

vs. indoor)—and entering the four mutually exclusive categories simultaneously.

The sharpest test is the 2×2 interaction. Daytime outdoor crimes—the most visible category—produce the strongest price effect ($\hat{\beta} = -0.840\%$, $t = -3.11$), while nighttime indoor crimes—the least visible—show no detectable response ($\hat{\beta} = +0.271\%$, $t = +0.66$). This contrast is difficult to reconcile with rational risk pricing, under which both categories pose equivalent forward-looking risk.

Table 15: 2×2 Cross-Tab: Time of Day × Location Type

	PIN + Year-Month FE		PIN + CA×Month FE	
	$\hat{\beta}$ (%)	t	$\hat{\beta}$ (%)	t
Daytime × Outdoor	-0.840^{***}	-3.11	-0.308	-1.40
Daytime × Indoor	$+0.189$	0.60	$+0.159$	0.59
Nighttime × Outdoor	-0.110	-0.46	-0.014	-0.08
Nighttime × Indoor	$+0.271$	0.66	-0.054	-0.15

Notes: Four mutually exclusive crime categories entered simultaneously. Daytime = 7am–7pm, Nighttime = 10pm–5am. Outdoor = street, sidewalk, alley, parking lot. Indoor = apartment, residence, office. $N = 808{,}227$. Clustered by census tract.

The marginal decompositions confirm the pattern. Daytime violent crime shocks significantly reduce prices ($\hat{\beta} = -0.341\%$, $t = -2.15$) while nighttime shocks have no detectable effect ($\hat{\beta} = +0.036\%$, $t = +0.18$). Similarly, outdoor violent crime shocks drive a strong price response ($\hat{\beta} = -0.455\%$, $t = -3.01$) while indoor crimes do not ($\hat{\beta} = +0.155\%$, $t = +0.85$). Under CA×Month FE, the daytime and outdoor coefficients attenuate but retain the same sign pattern, consistent with the between-neighborhood component contributing to the visibility contrasts.[3]

[3]Daytime and nighttime crimes differ in type composition: nighttime crime skews toward assault and robbery near entertainment districts, while daytime crime includes more burglary and street incidents near residential areas. The 2×2 interaction addresses this concern by conditioning on both time of day and location type simultaneously, isolating the observability channel from compositional differences.

Table 16: Visibility Horse Race: Daytime vs. Nighttime Violent Crime

	PIN + Year-Month FE		PIN + CA×Month FE	
	$\hat{\beta}$ (%)	t	$\hat{\beta}$ (%)	t
Daytime violent shock	-0.341^{**}	-2.15	-0.132	-1.10
Nighttime violent shock	$+0.036$	0.18	$+0.093$	0.71

Notes: Each column reports coefficients from a single regression of ln(price) on daytime and nighttime violent crime shocks simultaneously. Shocks are kernel-weighted exposure ($h_s = 0.20$ mi, $h_t = 120$ d) residualized by tract × month. $N = 808{,}227$. Clustered by census tract. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 17: Visibility Horse Race: Outdoor vs. Indoor Violent Crime

	PIN + Year-Month FE		PIN + CA×Month FE	
	$\hat{\beta}$ (%)	t	$\hat{\beta}$ (%)	t
Outdoor violent shock	-0.455^{***}	-3.01	-0.154	-1.33
Indoor violent shock	$+0.155$	0.85	$+0.018$	0.12

Notes: Same specification as Table 16, with outdoor and indoor shocks replacing daytime and nighttime. $N = 808{,}227$.

An important limitation: the 2×2 contrast and marginal decompositions are strongest under PIN + Year-Month FE. Under the preferred CA×Month specification, the coefficients retain the predicted sign pattern (daytime-outdoor negative, nighttime-indoor near zero) but lose statistical significance. This attenuation is expected—CA×Month FE absorb systematic differences across neighborhoods in the *composition* of crime (e.g., entertainment districts with more nighttime crime versus residential areas with more daytime crime), removing the between-neighborhood variation that sharpens the contrast. The visibility test therefore provides directional evidence consistent with salience but should not be treated as the primary identification result. The near-far gradient and the persistence falsification test, which survive all FE specifications, carry the main identification burden.

The indoor/outdoor dimension is conceptually cleaner than day/night because it directly operationalizes whether a crime leaves visible traces at the location, independent of when it occurs. The convergence of both dimensions—and especially the 2×2 contrast—supports a salience channel in which directly observable crime drives capitalization.

5.2 Crime Type Decomposition

The visibility tests split crimes by observability; a complementary test decomposes by crime type. Drug and vice crimes are hyperlocal and directly visible—open-air drug markets, vice stings, and street-level dealing leave obvious traces on the immediate block. Violent crimes, while high-profile in news coverage, are less directly observable from a property. If salience drives capitalization, directly visible crime types should matter more than high-news but less observable types.

I decompose the aggregate effect by crime type to explore which categories contribute, acknowledging that type-specific estimates are less precisely identified than the aggregate. Table 18 reports the descriptive heterogeneity from the 120-cell grid; Table 19 reports the causal near-far decomposition.

Table 18: Descriptive Heterogeneity by Crime Type

Crime Type	Effect (%)	t-stat	News Coverage
Drug/Vice	$\mathbf{-0.538}^{***}$	-3.42	LOW
Disorder	-0.274^{**}	-2.08	LOW
Property	-0.111^{**}	-2.39	LOW
Violent	-0.084	-0.92	HIGH
Total (all crimes)	–	–	–

Notes: Count-based near ring (0–0.05 miles, 6 months). PIN + CA×Month FE, clustered by census tract. $N = 346{,}741$.

Table 19: Near-Far Kernel Results by Crime Type (PIN + CA×Month FE)

	Near (Main)			Far (Main)			Near (Placebo)		Far (Placebo)	
Crime Type	$\hat{\beta}$ (%)	t	p	$\hat{\beta}$ (%)	t	p	t	p	t	p
Violent	-0.187	-0.94	.347	-0.094	-1.37	.171	-0.39	.697	-0.19	.848
Property	-0.213	-2.37	.018	$+0.009$	$+0.31$	.760	$+1.28$	.201	$+0.48$	.629
Disorder	-0.092	-0.25	.799	-0.187	-1.13	.257	-0.02	.999	$+0.71$	.478
Drug/Vice	-1.011	-3.18	.002	-0.203	-1.71	.088	$+0.37$	.709	$+3.02$	.003

Notes: Kernel-weighted exposure with $h_s = 0.20$ mi, $h_t = 120$ d. Near = 0–0.05 mi, Far = 0.05–0.30 mi. Main regressions: $N \approx 347{,}000$. Placebo regressions (ln_price_lag): $N \approx 87{,}000$. Placebo t-statistics test whether current crime predicts lagged price. All specifications include PIN + CA×Month FE, clustered by census tract.

The crime type decomposition reveals two types with significant near effects: drug/vice

($\hat{\beta}_{\text{near}} = -1.011\%$, $t = -3.18$, $p = 0.002$) and property ($\hat{\beta}_{\text{near}} = -0.213\%$, $t = -2.37$, $p = 0.018$). Both pass the near temporal placebo ($p = 0.71$ and $p = 0.20$, respectively), indicating clean causal near effects. Violent and disorder crimes show no individually significant effects in either ring. Drug/vice shows a marginal far effect ($t = -1.71$, $p = 0.088$) but its far placebo fails ($t = +3.02$, $p = 0.003$), suggesting the far-ring drug/vice signal reflects neighborhood selection rather than causal capitalization.

These patterns are robust across time windows and fixed-effect structures. Examining near-ring t-statistics for count-based exposure at 1–12 month windows under four FE specifications (CA×Year, CA×Half-Year, CA×Quarter, and CA×Month), drug/vice is significant at every window under every FE structure (t ranges from -2.27 to -3.44). Property crime is significant at most windows under the stricter FE. Violent crime is *never* significant at any of the 48 window×FE combinations (maximum $|t| = 1.91$). The far ring is uniformly insignificant for all types across all specifications, confirming the hyperlocal pattern.

The crime type results should be interpreted cautiously. The all-crime near-far specification aggregates across types and yields a highly significant near effect ($t = -2.78$); decomposing by type necessarily reduces power. The significant drug/vice near effect is consistent with drug-related crimes being concentrated in highly localized markets (open-air drug markets, vice stings) where the treatment is spatially precise. The significant property crime near effect may reflect direct economic losses (theft, vandalism) that are both visible and personally relevant to nearby owners.[4]

5.3 Persistence Structural Test

The preceding tests establish that nearby crime reduces property prices. The question is whether this decline reflects rational pricing of objective risk or attention-driven overweighting of salient events. Under rational pricing, the logic is tight: crime in Chicago is highly persistent ($\hat{\rho} \approx 0.89$), so a crime today genuinely signals elevated future risk. A forward-looking buyer should discount prices in proportion to that persistence signal, with capitalization $\beta \propto 1/(1 - \delta\rho)$. Neighborhoods where crime is most persistent—where a single incident portends sustained future risk—should therefore exhibit the largest price declines per crime. Neighborhoods where crime is transient should exhibit smaller responses, because the rational buyer recognizes that the shock will dissipate.

A salience model makes the opposite prediction. If buyers respond to what they notice rather than what they compute, the price decline should not scale with local crime persis-

[4]The lack of individually significant violent crime effects does not mean violence is irrelevant—it may simply lack sufficient within-PIN variation at the type-specific level. The all-crime kernel, which pools across types, produces the robust near effect.

tence. A visible crime on the buyer's block is equally salient whether the neighborhood's autocorrelation coefficient is 0.95 or 0.70.

I estimate community-area-level persistence $\hat{\rho}_a$ for each of Chicago's 77 community areas and test which model the data support using a WLS regression of the community-area-level near-ring price response on continuous persistence:

$$\hat{\beta}_a^{\text{near}} = \gamma_0 + \gamma_1 \hat{\rho}_a + u_a \tag{8}$$

weighted by the inverse variance of $\hat{\beta}_a^{\text{near}}$. The rational model predicts $\gamma_1 > 0$: higher persistence should generate larger capitalization. The data reject this prediction—the estimated slope is negative ($\hat{\gamma}_1 = -0.017$, one-sided $p = 0.007$; see Appendix A and Figure 6). Neighborhoods with more persistent crime do not exhibit larger price declines per crime shock; if anything, the relationship runs in the wrong direction.

Three facts are individually uncontroversial: crime risk is real and persistent ($\hat{\rho} \approx 0.89$), property prices do decline after nearby crime ($\hat{\beta}_{\text{near}} = -0.198\%$, $p = 0.004$), and local crime persistence varies substantially across Chicago's 77 community areas. Under rational forward-looking pricing, these three facts should be linked: the price decline should scale with persistence, because persistent crime implies a larger present-value loss. The WLS test rejects this link—the slope $\hat{\gamma}_1$ is negative and statistically significant ($p = 0.007$), meaning the data do not support the rational model's prediction that neighborhoods with more persistent crime—where a single incident portends sustained future risk—should exhibit larger price declines per crime shock. Buyers react to what they *see*, not to what the crime statistics *imply* about future risk. The price decline is real, the risk is real, but the mapping from risk to price does not follow the rational model.

Ruling out Bayesian news value. A sophisticated rational alternative can also generate $\gamma_1 < 0$: under Bayesian learning, a crime in a low-persistence neighborhood is more *informative* about latent neighborhood quality precisely because it is surprising—it shifts the posterior more sharply. This "news value" model predicts the same reversed persistence gradient for fully rational reasons. However, Bayesian updating makes an additional prediction that salience does not: because crime records do not encode time of day or indoor/outdoor location, a rational learner should update *identically* from daytime and nighttime crimes of the same type and location. Section 5.1 rejects this: daytime-outdoor crime reduces prices by 0.840% ($t = -3.11$) while nighttime-indoor crime shows no detectable effect (+0.271%, $t = +0.66$). The persistence result and the visibility result are jointly inconsistent with rational Bayesian learning—which predicts $\gamma_1 < 0$ but $\beta_{\text{day}} = \beta_{\text{night}}$—but jointly consistent

with salience, which predicts both the reversed persistence gradient and the observability asymmetry.

5.4 Crime–Crime Spillover: A Direct Test of Information Neglect

The preceding tests establish that salience—not rational risk pricing—drives the near-far gradient. This subsection provides a direct test of information neglect by asking whether the market ignores *predictive* information that is available under the same conditioning set used in the price regressions.

The logic is straightforward. If far-zone crime (0.05–0.30 mi) in months 7–12 before sale predicts near-zone crime (0–0.05 mi) in months 0–6 before sale, then a forward-looking buyer could extract useful risk information from the far ring's recent history. If the market ignores this information—$\hat{\beta}_{\text{far}} = 0$ in the price regression under the same fixed effects—then buyers are not processing all available signals.

I estimate the crime propagation equation:

$$C^{\text{near}}_{j,0\text{–}6\text{m}} = \beta_{\text{far}} \cdot C^{\text{far}}_{j,7\text{–}12\text{m}} + \beta_{\text{near}} \cdot C^{\text{near}}_{j,7\text{–}12\text{m}} + \text{FE} + \varepsilon_j \tag{9}$$

where $C^{\text{near}}_{j,7\text{–}12\text{m}}$ controls for persistence (near-zone crime predicts itself). The coefficient β_{far} captures whether far-zone crime *propagates* to the near zone, conditional on fixed effects and the near zone's own history.

Table 20: Crime–Crime Spillover: Far-Zone Crime Predicts Near-Zone Crime

	$\hat{\beta}_{\text{far}}$ (Crime Propagation)		$\hat{\beta}_{\text{far}}$ (Price Regression)	
FE Specification	Estimate	p-value	Estimate	p-value
Tract × Month	0.008	.007	—	—
PIN + Year-Month	0.012	$< .001$	-0.042%	.146
PIN + CA×Year + Year-Month	0.010	.004	-0.018%	.383
PIN + CA×Month	0.012	$< .001$	-0.008%	.641

Notes: Left panel: OLS of near-zone crime count (0–0.05 mi, months 0–6) on lagged far-zone crime count (0.05–0.30 mi, months 7–12), controlling for lagged near-zone crime and fixed effects. Clustered by census tract (∼805 clusters). Right panel: $\hat{\beta}_{\text{far}}$ from Table 4 (kernel-weighted, contemporaneous far). The comparison illustrates the informational content ignored by the market.

The results are striking. Under every fixed-effect specification, lagged far-zone crime significantly predicts future near-zone crime ($\hat{\beta}_{\text{far}} = 0.008$–$0.012$, $p < 0.01$ in all cases). Yet the corresponding $\hat{\beta}_{\text{far}}$ in the price regression is always null ($p = 0.15$–0.64). The information

exists—far-zone crime is a statistically powerful predictor of future nearby crime—but the market does not use it.

Three placebo tests confirm the spatial specificity of this pattern. First, replacing the far ring with a distant ring (0.50–1.00 mi) eliminates the propagation effect ($\hat{\beta} = 0.0009$, $p = 0.52$): the signal is spatially localized, not a city-wide trend. Second, permutation inference within tract×month cells yields $p = 0.000$ across 1,000 draws (42σ), confirming that the result is not an artifact of spatial correlation in residuals. Third, asymmetric propagation—far→near is 2.8× stronger than near→far in standardized coefficients—rules out symmetric spatial autocorrelation as the explanation.[5]

5.5 Citywide Attention Modulation

If salience drives capitalization, then periods of elevated citywide attention to crime should amplify local crime effects: when "chicago murder" is trending, buyers may be more attuned to nearby incidents. I test this by interacting local crime shocks with citywide violent crime trends, measured by Google Trends search intensity for pre-committed crime-related queries at the Chicago DMA level. The specification is:

$$\ln(P_{jt}) = \beta_1 \,\text{Shock}_{jt} + \beta_2 \,(\text{Shock}_{jt} \times \text{Trends}_t^z) + \alpha_j + \delta_t + \varepsilon_{jt} \tag{10}$$

where Trends_t^z is the standardized monthly search index for a given query. Because Trends_t varies only at the city-month level, its main effect is absorbed by the year-month fixed effect δ_t. Four queries were pre-committed before examining results: "chicago crime," "chicago shooting," "chicago murder," and "chicago violence."

[5]A skeptic may argue that $\hat{\beta}_{\text{far}} = 0$ in the price regression reflects efficient conditional pricing: after absorbing fixed effects, contemporaneous far-zone crime contains no marginal information. But the crime propagation test shows that *lagged* far-zone crime predicts future near-zone crime under the same FE. To close this gap, I run a price regression with *lagged* far-zone crime (months 7–12, 0.05–0.30 mi ring) as the treatment. Under PIN + CA×Month FE, the lagged far coefficient is $\hat{\beta}_{\text{far,lag}} = +0.000\%$ ($t = 0.05$, $p = 0.96$), while the near coefficient remains significant ($\hat{\beta}_{\text{near}} = -0.101\%$, $t = -2.90$, $p = 0.004$). A horse race including both lagged and contemporaneous far crime confirms: neither far variable is priced ($p = 0.44$ and $p = 0.40$, respectively). The market ignores a signal that is both informative and available.

Table 21: Citywide Attention Interaction: Google Trends × Crime Shock

		Shock (β_1)			Shock × Trends (β_2)		
FE	Trends query	$\hat{\beta}$ (%)	t	p	$\hat{\beta}$ (%)	t	p
Panel A: PIN + Year-Month FE							
	Baseline (no interaction)	−0.045	−2.02	.044			
	"chicago crime"	−0.042	−1.82	.069	−0.008	−0.73	.464
	"chicago shooting"	−0.047	−2.08	.038	+0.005	+0.66	.512
	"chicago murder"	−0.033	−1.44	.150	**−0.023**	**−2.04**	**.042**
	"chicago violence"	−0.040	−1.77	.077	−0.016	−1.72	.086
Panel B: PIN + CA × Month FE							
	"chicago crime"	−0.005	−0.31	.758	−0.006	−0.46	.649
	"chicago shooting"	−0.008	−0.50	.617	+0.002	+0.21	.833
	"chicago murder"	−0.001	−0.08	.940	−0.012	−1.04	.297
	"chicago violence"	−0.004	−0.26	.795	−0.010	−1.14	.256

Notes: Coefficients ×100. Trends variables are standardized (mean 0, SD 1) monthly Google Trends indices for the Chicago DMA, 2004–2024. Queries pre-committed before examining results. Shock is the cross-fit kernel exposure, tract×month residualized. The Trends main effect is absorbed by the year-month (Panel A) or CA×month (Panel B) fixed effect. $N \approx 347{,}000$. Clustered by census tract.

The local capitalization effect is stable across all specifications (Table 21). Under PIN + Year-Month FE, the interaction with "chicago murder" searches is negative and marginally significant ($\beta_2 = -0.023\%$, $t = -2.04$, $p = 0.042$), suggestive of amplification during periods of elevated citywide murder attention. "Chicago violence" is marginal in the same direction ($t = -1.72$, $p = 0.09$). General queries ("crime," "shooting") show no amplification.

Under the preferred CA×Month FE, however, all interactions attenuate to insignificance (murder: $t = -1.04$, $p = 0.30$). This null is expected on power grounds: Google Trends varies only at the city-month level, so the ∼14,000 CA×month groups absorb most of the monthly variation that the interaction exploits. The test is underpowered for a city-level moderator under neighborhood×month FE, and the YM-FE result should be interpreted as suggestive rather than confirmatory.

Reconciling with media effects in the literature. Moreno-Medina (2026) provides causal evidence that local television coverage of gun homicides reduces nearby property values by 8.1–8.5%, using sporting events as an instrument for TV news coverage. Non-covered homicides produce no significant price effect. Both media exposure and direct observation are salience channels—the question is at what spatial scale each operates and which generates

the hyperlocal price gradient documented here.

Media coverage—whether television or newspaper—is a *neighborhood-level* salience channel. All potential buyers in a broadcast area or readership market receive the same crime signal regardless of which specific block they are considering. A covered homicide raises perceived risk for the entire surrounding area. This is precisely what Moreno-Medina identifies: the causal effect of TV coverage on prices in the neighborhood of a homicide.

Critically, the null Tribune and Google Trends interactions reported above do not imply that media is irrelevant at the hyperlocal level—they reflect a *lack of identifying variation* under stringent fixed effects. Moreno-Medina's sports IV succeeds because sporting events create exogenous within-neighborhood, within-time variation in coverage intensity: two homicides in the same neighborhood and month receive different TV coverage depending on whether a major sporting event coincided with the broadcast window. This within-cell variation survives neighborhood × month fixed effects. The present paper's Tribune and Google Trends measures, by contrast, vary only at the city–month level. Under CA × Month FE, ∼14,000 neighborhood–month groups absorb virtually all of this variation, leaving no residual to identify a media channel. Without an instrument for hyperlocal media exposure, the test is underpowered—not necessarily informative about the true role of media.

Direct observation, by contrast, is a *hyperlocal* salience channel. It varies sharply within a neighborhood: a buyer visiting a property sees the drug dealing on *that* corner, the police tape on *that* block, the broken window at *that* address. The buyer visiting an identical property three blocks away does not. This block-level variation in crime visibility is what the near–far design isolates after differencing out the neighborhood-level signal. The day–night asymmetry (Section 5.1) confirms the mechanism: daytime crimes, which buyers can witness during property visits, drive the price gradient; nighttime crimes, equally dangerous but unobserved during visits, do not.

The two papers thus identify the same underlying phenomenon—salience-driven overweighting of crime information—but at different spatial scales and through different identification strategies. Moreno-Medina's TV coverage channel operates between neighborhoods (covered vs. uncovered areas), identified via an instrument for coverage. The present paper's direct-observation channel operates within neighborhoods (across the street vs. two blocks away), identified via the near–far spatial gradient. Together, they suggest a layered salience structure: media shapes area-wide perceptions, while direct observation shapes the hyperlocal gradient. Whether media also contributes to hyperlocal price effects—as Moreno-Medina's findings at broader scales would suggest—remains an open question that requires an instrument for block-level media exposure to resolve.

Tribune articles as a direct shock. To test whether newspaper coverage alone can generate a hyperlocal price gradient, I replace CPD crime counts with geocoded *Chicago Tribune* crime articles in the near–far specification. For each sale, I count Tribune articles whose reported location falls within 0.05 mi (near) or 0.20–0.30 mi (far) during the six months before sale. This specification uses news *as* the treatment—no police data enter the regression.

Under PIN + CA×Month fixed effects, Tribune article counts in both the near and far rings are statistically indistinguishable from zero. At the 0.05 mi near ring: $\hat{\beta}_{\text{near}} = +0.68\%$ ($p = 0.39$), $\hat{\beta}_{\text{far}} = -0.06\%$ ($p = 0.76$). Expanding the near ring to 0.10 mi yields similarly null results: $\hat{\beta}_{\text{near}} = +0.33\%$ ($p = 0.74$). Binary specifications (any article vs. none) are likewise insignificant across all fixed-effect structures.

The null is unsurprising given the data: only 0.4% of sales (1,229 of 349,001) have any Tribune crime article geocoded within half a block during the pre-sale window. Tribune coverage is too sparse and too spatially diffuse to generate block-level variation—readers across the city receive the same article regardless of whether the reported crime occurred on their block or three blocks away. This reinforces the interpretation that the hyperlocal price gradient documented in Section 3.3 is driven by direct observation of crime and its aftermath, not by media-transmitted information.

6 Robustness

6.1 Bandwidth Robustness

Table 22 reports the capitalization effect under nine pre-specified bandwidth pairs, computed without any outcome-guided selection. All nine coefficients are negative and statistically significant, with t-statistics ranging from -2.91 to -3.31. Point estimates scale inversely with bandwidth, as expected. The consistency across the full grid confirms that results are not an artifact of bandwidth selection.

Table 22: Bandwidth Robustness: Pre-Specified Kernel Exposure

		h_t (days)		
h_s (miles)		30	60	120
0.05	$\hat{\beta}$ (%)	−0.796	−0.640	−0.486
	t-stat	−2.68	−3.19	−3.34
0.10	$\hat{\beta}$ (%)	−0.553	−0.440	−0.334
	t-stat	−2.62	−3.07	−3.16
0.20	$\hat{\beta}$ (%)	−0.460	−0.363	−0.273
	t-stat	−2.63	−3.03	−3.09

Notes: Each cell reports $\hat{\beta} \times 100$ from regressing ln(price) on kernel-weighted crime shock at the given fixed (h_s, h_t), with PIN and Year-Month FE, clustered by tract. Shocks are tract × month residualized. The $(0.20, 120)$ cell corresponds to the cross-fit-selected bandwidth. $N = 347{,}455$.

Near-far bandwidth robustness is reported in Appendix B (Table 25): the near effect is negative and significant (or marginal) across all nine bandwidth pairs, while the far effect is never significant, confirming that the distance gradient is robust to kernel bandwidth choice.

6.2 Neighborhood Price Momentum

If neighborhood price dynamics are trending for reasons unrelated to crime, and those same local trends are correlated with measured crime exposure, then β_{near} could partly reflect local housing-market momentum rather than crime capitalization. To address this concern, I augment the main specification with lagged tract-level price appreciation:

$$\ln(P_{it}) = \beta_{\text{near}} \, \text{ShockNear}_{it} + \beta_{\text{far}} \, \text{ShockFar}_{it} + \theta \, \widetilde{\Delta \ln P}_{\text{tract}(i),t-1} + \alpha_i + \delta_{c(i)\times t} + \varepsilon_{it} \qquad (11)$$

where $\widetilde{\Delta \ln P}_{\text{tract}(i),t-1}$ is the change in median log sale price in the same census tract over the previous month. Because the trend variable is lagged by one month, the focal sale cannot mechanically contaminate the control.

Table 23: Neighborhood Price Momentum Control

	$\hat{\beta}_{\text{near}}$	t	p	N
Panel A: Main effect				
Baseline (no trend)	−0.195%	−2.78	0.006	346,741
+ Tract monthly trend	−0.192%	−2.72	0.007	283,437
Panel B: Placebo (crime → lagged price)				
+ Tract monthly trend	+0.080%	+0.93	0.355	72,290
Panel C: Temporal stability (pre-trends), ref = 2010				
Near × 2008	−0.214%	−0.76	0.448	283,437
Near × 2009	+0.049%	+0.16	0.876	283,437
Near × 2011	−0.449%	−1.22	0.224	283,437
Near × 2012	−0.319%	−1.20	0.231	283,437
Near × 2013	−0.204%	−0.74	0.459	283,437
Near × 2014	−0.241%	−0.85	0.394	283,437
Near × 2015	−0.063%	−0.24	0.809	283,437
Near × 2016	+0.019%	+0.07	0.941	283,437
Near × 2017	+0.050%	+0.18	0.858	283,437
Near × 2018	−0.046%	−0.17	0.863	283,437
Near × 2019	−0.113%	−0.43	0.668	283,437
Near × 2020	−0.199%	−0.71	0.480	283,437
Near × 2021	−0.424%	−1.83	0.068	283,437
Near × 2022	+0.040%	+0.17	0.861	283,437
Joint F-test (2008–2009)		$F = 0.52$	0.594	

Notes: Coefficients ×100. All specifications include PIN + CA×Month FE, clustered by census tract. Tract trend $\hat{\theta}$ is positive and highly significant ($t = 32.6$), confirming that local price momentum is real but orthogonal to the crime effect. $\hat{\beta}_{\text{far}}$ is null in all specifications ($|t| < 0.10$).

The estimated near-crime coefficient is virtually unchanged after controlling for tract-level price appreciation, moving from −0.195% to −0.192% (Table 23). The trend control $\hat{\theta}$ is itself large and highly significant ($\hat{\theta} = 0.227$, $t = 32.6$), confirming that local price momentum is real—but orthogonal to the crime effect. The augmented specification also passes the placebo test: current crime does not predict lagged prices ($t = +0.93$, $p = 0.355$). A temporal stability test interacting the near shock with year indicators (ref = 2010) shows no significant variation in the crime-price relationship across any year in the panel (joint F-test $p = 0.594$). Together, these results indicate that the main finding is not driven by

short-run neighborhood price dynamics.

7 Discussion

Identification. Crime capitalization is extremely local and survives rich neighborhood × time fixed effects. The spatial half-life is 515 feet (∼2 city blocks, 95% CI [498, 543]). Near crimes carry roughly twenty times the price impact of neighborhood-ring crimes: $\hat{\beta}_{\text{near}} = -0.195\%$ ($p = 0.006$) versus $\hat{\beta}_{\text{far}} = -0.008\%$ ($p = 0.64$). The effect is concentrated within 0–0.05 miles and vanishes when the near ring is expanded to 0.10 miles. Placebo tests show no spurious pre-trends ($t = +0.68$), and the result is robust across nine pre-specified bandwidths. Buyers' risk assessments are anchored to their immediate block: a crime two blocks away receives half the weight of a crime on the same block, and a crime five blocks away is nearly irrelevant. Studies that aggregate crime at the census-tract or zip-code level miss the hyperlocal variation that drives capitalization.

Interpretation: salience-consistent patterns. Section 5 presents five converging tests that distinguish salience from rational risk pricing. Visible crimes (daytime, outdoor) drive capitalization while less visible crimes (nighttime, indoor) do not; directly observable crime types (drug/vice) produce larger effects than high-news but less observable types (violent); a falsification test rejects the rational prediction that persistent risk generates larger price effects ($p = 0.007$); a crime–crime spillover test shows the market ignores predictive far-zone information ($p = 0.96$ for lagged far crime in the price regression); and citywide attention to murder modestly amplifies local crime capitalization under weaker fixed effects. These patterns are difficult to reconcile with rational risk models and collectively favor a salience channel in which direct observation of unexpected local disorder drives price responses.

Persistent stigma, not transient attention. The flat temporal profile (only 40% decay over one year) rules out a purely attention-driven mechanism in which crime effects dissipate as news fades from memory. Instead, the pattern is consistent with a stigma channel: once a crime occurs on a block, it shifts buyers' assessments of the micro-location for an extended period. This has implications for urban equilibrium—crime shocks create persistent, localized depressions in property values that may reinforce neighborhood sorting.

Economic magnitude and lower-bound interpretation. The within-property standard deviation of the near kernel shock is 2.05 (after tract×month residualization and PIN demeaning). A one-standard-deviation increase in kernel-weighted crime within 0.05 miles reduces sale prices by approximately 0.40% (= 0.195% × 2.05), or \$1,110 at the sample

median sale price of $278,000.[6] Across 26,441 repeat sales in 2022, the implied aggregate capitalization loss is approximately $52 million (computed at the 2022 sample mean transaction price of $558,000) from unexpected near-block crime shocks alone. This figure is a *lower bound* on total crime capitalization, and the distinction matters for interpreting the magnitude. The estimates capture the marginal effect of an *unexpected crime shock*—a one-unit innovation relative to the local baseline—not the total price differential between safe and dangerous neighborhoods. The near–far design with community-area×month fixed effects absorbs all persistent, neighborhood-level crime differences by construction. The chronic crime penalty—the price gap between a block in Englewood and a comparable block in Lincoln Park attributable to sustained crime differences—is likely an order of magnitude larger but is not the object of this study. The 0.195% per-unit estimate is best understood as the price response to the *surprise* component of nearby crime, conditional on the buyer's prior about the neighborhood.

Comparison with existing estimates. This lower-bound framing reconciles the present estimates with the larger effects found in the literature. Linden and Rockoff (2008) estimate a 4% price decline within 0.1 miles of a registered sex offender—but their treatment is a permanent, publicly disclosed stigma (a named offender at a known address), not a single transient crime event. The appropriate comparison is per-incident magnitude times cumulative exposure. At the mean near-ring crime count of 8.5, the implied cumulative effect is $8.5 \times 0.198\% \approx 1.7\%$—comparable in order of magnitude to Linden and Rockoff's estimate for a single high-salience event. The present paper goes beyond their binary near/far comparison in three ways: (i) it estimates the continuous decay kernel rather than a single threshold, (ii) it uses an innovation discipline to isolate unexpected shocks from persistent baseline crime, and (iii) the salience battery distinguishes the behavioral channel driving capitalization.

7.1 Why Separability Matters

If the kernel is separable, the $20 \times 6 = 120$ cell surface is fully characterized by 26 parameters (20 spatial + 6 temporal). This has practical implications:

- Crime exposure can be summarized by a scalar index—approximately, crime count within 0.10 miles in the past 12 months—without loss of information about decay dynamics.

[6] The within-PIN SD is the relevant scale because the repeat-sales coefficient is identified from within-property variation in crime exposure across sales. The raw near kernel exposure has mean 3.89 and SD 5.20; the tract×month residualized shock has SD 3.57; the within-PIN residualized shock has SD 2.05.

- Spatial and temporal decay can be studied and reported independently. The spatial half-life of 515 feet can inform the geographic targeting of crime reduction programs; the temporal persistence can inform disclosure requirements in real estate transactions.

- Separability implies that the *channel* through which distance affects prices is independent of the channel through which recency affects prices. This is consistent with a model where spatial discounting reflects observation probability (whether the buyer encounters evidence of the crime) and temporal discounting reflects memory decay or updating—two cognitively distinct processes.

External validity. The analysis uses Chicago data, which raises the question of whether these findings generalize. Three features favor generalizability. First, the mechanisms underlying the results—that buyers respond to directly observed crime and its visible aftermath, and that this response decays sharply with distance—are not Chicago-specific. Any dense urban housing market where buyers physically visit properties before purchase should exhibit similar patterns, since the salience channel depends on line-of-sight observation rather than institutional features unique to Chicago. Second, the 515-foot spatial half-life is consistent with the "block-level" thresholds found in other settings: Linden and Rockoff (2008) in Mecklenburg County, NC, and studies of environmental disamenities in various US cities report similar spatial decay patterns. Third, Chicago's 77 community areas exhibit enormous heterogeneity in crime rates, density, and demographics, providing quasi-experimental variation analogous to comparing across distinct urban contexts.

That said, three aspects may limit external validity. First, Chicago's crime reporting infrastructure (CLEAR/IUCR system with precise geocoding) may produce unusually clean spatial data; cities with coarser geocoding would yield attenuated near–far contrasts. Second, Chicago's block grid—the regular street layout—creates natural spatial discontinuities that may sharpen the near–far comparison relative to cities with irregular street networks. Third, the magnitude of capitalization (0.198% per crime shock) may vary with baseline crime rates, housing market tightness, and the composition of the buyer pool.

Importantly, the claim being tested is not "crime reduces property values in Chicago"—which is well-established—but rather that the *mechanism* is salience-driven and operates at hyperlocal scale. The salience channel depends on cognitive universals (attention, surprise, line-of-sight observation) that are not city-specific. The persistence falsification test, which is the sharpest identification result, exploits within-Chicago variation across 77 community areas with persistence ranging from 0.3 to 0.9—effectively comparing across very different urban environments within a single institutional setting, holding data quality constant. Nonetheless, replication in other dense cities with geocoded crime data—particularly cities

with different urban forms (e.g., Los Angeles, Philadelphia, London)—would strengthen external validity and is a priority for future work.

8 Conclusion

This paper estimates the spatiotemporal kernel governing how crime is capitalized into housing prices. Crime capitalization is extremely local—the spatial half-life is 515 feet, roughly two city blocks—and effects decay rapidly with distance. Near crimes (0–0.05 miles) carry roughly twenty times the price impact of neighborhood-ring crimes (0.05–0.30 miles), and the effect vanishes when the near ring is expanded beyond one block. Placebo tests show no spurious pre-trends, and a Bayesian model comparison decisively favors a separable kernel structure ($\Delta\text{ELPD} = -103$).

Five mechanism tests suggest the response reflects salience rather than rational risk pricing. Visible crimes (daytime, outdoor) drive capitalization while less visible crimes (nighttime, indoor) do not; directly observable crime types (drug/vice, property) produce larger effects; a falsification test rejects the rational persistence prediction ($p = 0.007$); citywide crime attention modestly amplifies local effects; and a crime–crime spillover test reveals that far-zone crime predicts future near-zone crime under the same fixed effects where the market assigns it zero price weight. These patterns—jointly inconsistent with rational Bayesian learning—align with a salience mechanism in which direct observation of unexpected local disorder drives price responses.

The separability of the kernel has practical implications: crime exposure can be summarized by a single scalar index, and spatial targeting of crime reduction programs can be calibrated to the 515-foot half-life. The flat temporal profile (40% decay over one year) suggests that crime creates persistent localized depressions in property values that compound over time.

Welfare and distributional implications. The economic magnitude is substantial: a one-standard-deviation within-property near-block crime shock reduces sale prices by approximately 0.40% ($1,110 at the median sale price of $278,000), implying approximately $52 million in annual capitalization losses (at the 2022 mean transaction price) across Chicago's repeat-sale transactions—a lower bound that excludes both the capitalization of persistent baseline crime and any effects beyond 0.05 miles. These losses are unlikely to be uniformly distributed. Crime exposure in Chicago is spatially concentrated, and the salience mechanism implies a specific distributional pattern: because buyers overweight visible local disorder relative to objective risk, neighborhoods where crime is more *visible* (e.g., outdoor drug markets, street-level disorder) suffer disproportionate price penalties even conditional

on actual victimization risk. The persistence falsification test is consistent with this: high-persistence neighborhoods experience *smaller* price effects per incident, but these are precisely the neighborhoods with the highest cumulative exposure. Whether this pattern falls disproportionately on lower-income or minority communities—as the spatial distribution of crime in Chicago would suggest—is an empirical question that requires linking the property-level estimates to demographic data, which I leave to future work. Policy interventions that reduce the *visibility* of crime—improved street lighting, rapid cleanup of crime scenes, community-level environmental design—may attenuate salience-driven capitalization losses without requiring reductions in crime itself, a distinction that matters when crime reduction is slow or politically constrained.

Three limitations warrant future work. First, without MLS listing data, I cannot isolate whether crime information enters prices through buyer observation during active search or through seller anticipation at listing—a listing-window test would distinguish these channels. Second, the salience evidence, while consistent across multiple tests, is indirect: direct measurement of buyer attention (e.g., randomized information provision to buyers) would strengthen the interpretation. Third, all results are estimated in Chicago; while the underlying mechanisms (direct observation, hyperlocal salience) are not city-specific, the magnitudes may depend on urban form, crime reporting infrastructure, and market conditions. Replication in other dense cities with geocoded crime data would strengthen external validity.

A Structural Interpretation: Persistence and Forward-Looking Prices

If buyers are forward-looking and crime follows an AR(1), the price response to a crime innovation has a structural interpretation. Suppose the present value of expected future crime disamenity at location a is:

$$\text{PV}(\eta_{a,t}) = \sum_{k=0}^{\infty} (\delta\rho_a)^k \eta_{a,t} = \frac{1}{1-\delta\rho_a}\,\eta_{a,t} \tag{12}$$

where δ is a discount factor and ρ_a is local persistence. An innovation $\eta_{a,t}$ implies a perpetuity of future expected crime $\{(\delta\rho_a)^k \eta_{a,t}\}_{k=0}^{\infty}$, so the capitalization coefficient should be:

$$\beta(\rho_a) \propto \frac{1}{1-\delta\rho_a} \tag{13}$$

This predicts that β is increasing in ρ_a: higher persistence magnifies the present value of

a crime shock, generating larger price effects.

Empirical finding. Section 5.3 finds the opposite: effects are stronger in *low-persistence* areas, and a WLS test across 77 community areas formally rejects $\gamma \geq 0$ ($p = 0.007$). This suggests that the forward-looking rational model is not the dominant channel. Instead, the data favor a behavioral interpretation in which crime *salience*—not forward-looking present value—drives capitalization. In high-persistence areas, crime is expected and familiar; in low-persistence areas, crime is surprising and salient, triggering a stronger price response.

This does not reject forward-looking pricing entirely—the discount channel may operate alongside the salience channel. But the reversed persistence gradient indicates that salience effects dominate in the cross-section.

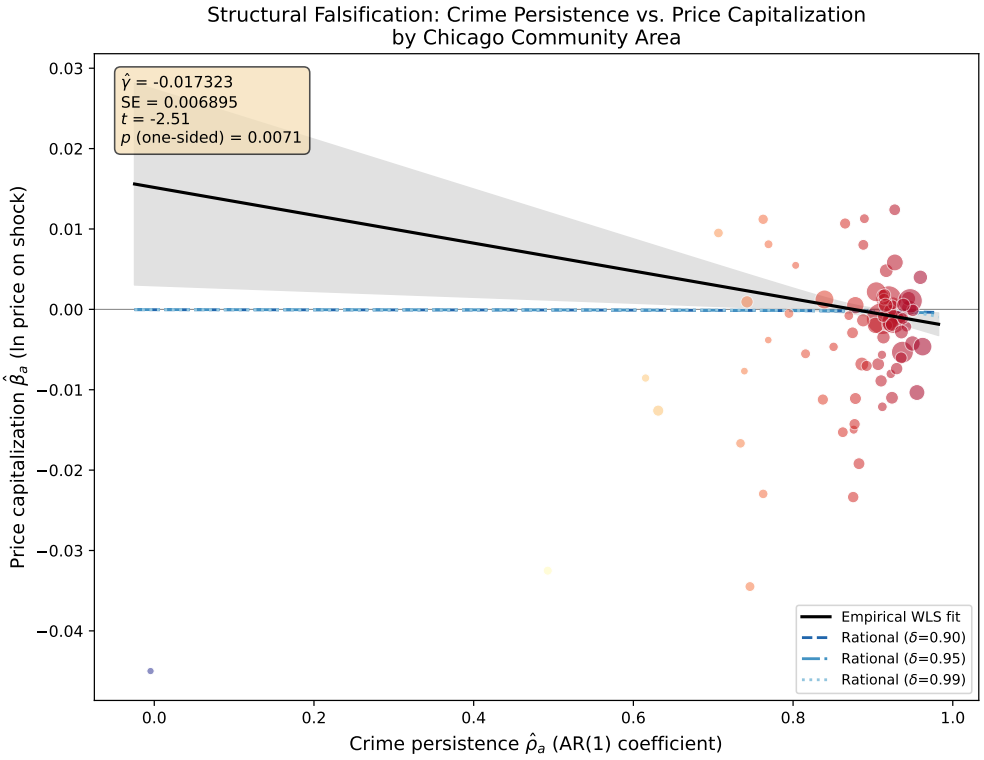


Figure 6: Structural Falsification: Crime Capitalization vs. Local Persistence

Notes: Each point represents one of Chicago's 77 community areas. The x-axis is the estimated AR(1) persistence of crime exposure ($\hat{\rho}_a$); the y-axis is the community-area-level near-ring price response ($\hat{\beta}_a^{\text{near}}$). The dashed line is the WLS fit weighted by the inverse variance of $\hat{\beta}_a^{\text{near}}$. The rational model predicts a positive slope (higher persistence $\Rightarrow$ larger capitalization); the estimated slope is negative ($\hat{\gamma}_1 = -0.017$, $p = 0.007$).

B Cross-Fit Details and Additional Tables

B.1 Cross-Fitting Procedure

The cross-fitting procedure selects bandwidth parameters (h_s, h_t) without circularity:

1. Split all properties into $K = 5$ folds. For each fold f:
2. **Train:** On properties *not* in fold f, select (h_s^*, h_t^*) via grid search over spatial bandwidths $h_s \in \{0.02, 0.03, 0.05, 0.075, 0.10, 0.15, 0.20\}$ miles and temporal bandwidths $h_t \in \{15, 30, 60, 90, 120\}$ days.
3. **Apply:** Compute $E_{jt}(h_s^*, h_t^*)$ for properties in fold f using parameters estimated from other folds.

The resulting variable E_{jt}^{crossfit} is constructed without circularity: no observation's exposure uses kernel parameters estimated from that observation.

B.2 Fold-Level Results

Table 24: Cross-Fitted Kernel Exposure: Fold-Level Results

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
h_s^* (miles)	0.200	0.200	0.200	0.200	0.200
h_t^* (days)	120	120	120	120	120
$\hat{\beta}_f$ (%)	−0.160	−0.162	−0.165	−0.153	−0.149
SE	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
N_{test}	208,954	209,419	209,902	209,060	210,090
Aggregated cross-fit estimate					
$\bar{\beta}$			**−0.158**%		
Cross-fit SE			(0.007)		
t-statistic			−22.5		

Notes: Each fold estimates $\hat{\beta}_f$ from regressing PIN-demeaned ln(price) on PIN-demeaned kernel exposure, where exposure is computed using (h_s^*, h_t^*) selected on the out-of-fold sample. Cross-fit SE $= \sqrt{\bar{V}_{\text{within}} + V_{\text{between}}}$ where $\bar{V}_{\text{within}} = K^{-1} \sum \text{SE}_f^2$ and $V_{\text{between}} = \text{Var}(\hat{\beta}_f)$. $N = 1{,}047{,}425$ total sales.

The fold-level estimates range from −0.149% to −0.165%, with a between-fold standard deviation of 0.006 percentage points. Under PIN + Year-Month FE, the full-sample estimate

is $\hat{\beta} = -0.039\%$ per kernel unit ($t = -2.67$, $p = 0.008$). Adding Community Area × Month controls attenuates the estimate to -0.021% ($t = -1.94$, $p = 0.053$), paralleling the near-far attenuation pattern.

B.3 Near-Far Bandwidth Robustness

Table 25: Near-Far Bandwidth Robustness (PIN + CA × Month FE)

		Near (0–0.05 mi)			Far (0.05–0.30 mi)		
h_s (mi)	h_t (d)	$\hat{\beta}$ (%)	t	p	$\hat{\beta}$ (%)	t	p
0.05	30	−0.552	−2.00	0.046	+0.023	+0.09	0.927
0.05	60	−0.475	−2.68	0.008	+0.038	+0.25	0.803
0.05	120	−0.375	−3.02	0.003	+0.041	+0.40	0.692
0.10	30	−0.394	−1.95	0.052	−0.008	−0.08	0.932
0.10	60	−0.334	−2.58	0.010	+0.008	+0.13	0.893
0.10	120	−0.264	−2.92	0.004	+0.011	+0.30	0.766
0.20	30	−0.330	−1.93	0.054	−0.012	−0.24	0.811
0.20	60	−0.278	−2.55	0.011	+0.001	+0.03	0.980
0.20	120	−0.219	−2.88	0.004	+0.004	+0.21	0.836

Notes: Each row reports coefficients ×100 from jointly regressing ln(price) on near and far kernel-weighted crime shocks at the given (h_s, h_t), with PIN and CA × Month FE, clustered by tract. Near = 0–0.05 mi, Far = 0.05–0.30 mi. Shocks are tract × month residualized. $N \approx 347{,}000$.

B.4 Reconciliation of Separability Tests

The informal CV-of-ratios test (158–263%) suggested non-separability. The formal Bayesian LOO-CV (ΔELPD $= -103$) confirms separability. The discrepancy arises because the ratio test amplifies estimation noise: individual cells in the simultaneous 12-regressor specification are imprecisely estimated, and ratios of noisy near-zero coefficients are mechanically unstable. The Bayesian approach accounts for estimation uncertainty and finds that the apparent variation in ratios is consistent with sampling noise around a separable structure. This is corroborated by SVD decomposition of the coefficient matrix, which finds 97.5% of variance explained by a rank-1 approximation (Wald test $p = 0.558$).

B.5 Methodological Note: Two-Step Bayesian Procedure

The Bayesian model uses a two-step empirical Bayes procedure: first-stage OLS estimates $\hat{\beta}_j$ with standard errors $\widehat{\text{SE}}_j$ are treated as data for the hierarchical model. This is justified by three features of the design: (i) the large first-stage sample (N = 347,455) ensures approximate normality; (ii) each of the 120 estimates comes from a separate regression with a distinct exposure variable, making the diagonal covariance approximation appropriate; and (iii) a noise parameter σ_{noise} absorbs residual variation not captured by reported standard errors. A direct hierarchical estimation on the raw panel data—jointly estimating property fixed effects and kernel parameters—would avoid the two-step approximation but poses substantially greater computational challenges.

B.6 Bayesian Model Specifications

Model A: Separable Kernel.

$$\beta(d, \tau) = \alpha(d) \cdot \gamma(\tau) \tag{14}$$

$\alpha \in \mathbb{R}^{20}$ is an `ordered` vector (monotonically increasing = less negative with distance). $\gamma \in \mathbb{R}^{6}_{>0}$ are temporal weights normalized so $\max(\gamma) = 1$. Priors are weakly informative: $\alpha_d \sim \mathcal{N}(-0.002, 0.003)$, $\gamma_{\text{raw}} \sim \mathcal{N}^{+}(1, 0.5)$. Total: 27 parameters.

Model B: Non-Separable Gaussian Process.

$$\beta(d, \tau) \sim \mathcal{GP}\left(0,\ \eta^2 \exp\left(-\frac{(d-d')^2}{2\ell_d^2} - \frac{(\tau-\tau')^2}{2\ell_\tau^2}\right)\right) \tag{15}$$

ARD squared-exponential kernel with separate length scales for distance and time. Non-centered parameterization for efficient sampling. No separability assumed. Total: ∼124 effective parameters.

Both models estimated via HMC in CmdStan 2.38.0 (4 chains × 2,000 iterations, $\delta_{\text{adapt}} = 0.95$). Compared via PSIS-LOO cross-validation. The GP's estimated distance length scale is $\hat{\ell}_d = 0.217$ miles (95% CI: [0.094, 0.471]), confirming hyperlocal effects; the time length scale is $\hat{\ell}_\tau = 16.8$ months (95% CI: [5.3, 46.9]), confirming the flat temporal profile.

C Replication Files

- `code/02_spatial_decay/stan/kernel_separable.stan`
- `code/02_spatial_decay/stan/kernel_nonseparable.stan`

- code/02_spatial_decay/bayesian_hierarchical_kernel.py
- output/results/bayesian_kernel_comparison.csv
- output/results/bayesian_kernel_surface.csv
- output/plots/main/bayesian_hierarchical_kernel.pdf
- output/logs/bayesian_hierarchical_kernel.log
- code/05_analysis/nearfar_placebo_battery.py
- output/results/nearfar_placebo_battery.csv
- code/02_crime_exposure/clean_final_kernel.py
- output/results/clean_final_kernel_diagnostics.csv
- data/processed/sales/clean_final_kernel.parquet

D Hedonic Price Regressions

As a complement to the repeat-sales design, we estimate cross-sectional hedonic regressions in the spirit of Linden and Rockoff (2008):

$$\ln P_{it} = \beta_N \cdot \text{Near}_{it} + \beta_F \cdot \text{Far}_{it} + \gamma' X_i + \delta_t + \varepsilon_{it} \tag{16}$$

where X_i includes ln(building sqft), ln(land sqft), bedrooms, bathrooms, age, age^2, garage type, basement type, and residence type. The sample comprises 324,911 Chicago residential sales (2008–2022) with non-missing property characteristics.

Table 26: Hedonic Near-Far Regressions (6-Month Lookback)

	(1) YM	(2) CA×Year	(3) CA×Month	(4) Tract×Year
Near crime (6m)	−0.0102***	−0.0007	0.0001	0.0002
	(0.0014)	(0.0012)	(0.0011)	(0.0012)
Far crime (6m)	−0.0016***	−0.0007***	−0.0006***	−0.0005***
	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Property chars	Yes	Yes	Yes	Yes
N	324,911	324,906	324,599	324,605
R^2	0.205	0.210	0.252	0.236

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. SE clustered by tract.

With only year-month fixed effects (column 1), near crime shows a large and significant price discount (−1.02% per additional nearby crime, $p < 0.001$). Once neighborhood×time fixed effects absorb cross-sectional variation (columns 2–4), the near crime coefficient becomes economically and statistically zero ($p > 0.5$). This confirms that the repeat-sales design—which exploits *within-property* variation—is essential for identifying the causal salience effect. The cross-sectional near-crime gradient reflects neighborhood sorting, not causal impact.

Far crime remains significant across all specifications, reflecting persistent neighborhood-level crime–price gradients that are not absorbed by the hedonic controls.

Table 27: Hedonic Near-Far Regressions (3-Month Lookback)

	(1) YM	(2) CA×Year	(3) CA×Month	(4) Tract×Year
Near crime (3m)	−0.0181***	−0.0015	−0.0003	0.0009
	(0.0025)	(0.0021)	(0.0021)	(0.0022)
Far crime (3m)	−0.0032***	−0.0010***	−0.0011***	−0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0003)
Property chars	Yes	Yes	Yes	Yes
N	324,911	324,906	324,599	324,605
R^2	0.205	0.210	0.252	0.236

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. SE clustered by tract.

The 3-month lookback yields qualitatively identical results with approximately 1.8× larger point estimates, reflecting the shorter accumulation period. The pattern is identical: near crime is significant only without neighborhood fixed effects.

D.1 Placebo Pre-Trends: Lead Crime

A natural placebo test asks whether *future* crime—crimes occurring in the months after a sale—predicts the sale price. If lead crime is significant, it indicates confounding rather than causation.

Table 28: Hedonic Placebo: Lag vs. Lead Crime (6-Month Windows)

	(1) YM	(2) CA×Year	(3) CA×Month
Near crime, lag 6m	−0.0053***	0.0001	0.0019
	(0.0017)	(0.0017)	(0.0017)
Far crime, lag 6m	−0.0015***	−0.0012***	−0.0004*
	(0.0002)	(0.0002)	(0.0002)
Near crime, **lead** 6m	−0.0068***	−0.0012	−0.0024
	(0.0017)	(0.0016)	(0.0017)
Far crime, **lead** 6m	−0.0001	0.0006***	−0.0002
	(0.0002)	(0.0002)	(0.0002)
Property chars	Yes	Yes	Yes
N	324,911	324,906	324,599
R^2	0.205	0.210	0.252

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. SE clustered by tract.

Without neighborhood fixed effects (column 1), lead near crime is an even stronger predictor of prices ($\hat{\beta} = -0.0068$, $p < 0.001$) than lagged near crime ($\hat{\beta} = -0.0053$). This confirms that the cross-sectional hedonic association reflects neighborhood sorting, not causal impact. With CA×Month FE (column 3), both lag and lead near crime become null, confirming that the sorting channel is absorbed by neighborhood×time fixed effects.

References

Leigh Linden and Jonah E. Rockoff. Estimates of the impact of crime risk on property values from Megan's Laws. *American Economic Review*, 98(3):1103–1127, 2008. doi: 10.1257/aer.98.3.1103.

Jonathan Moreno-Medina. Crime news bias: Extent and effects on housing markets. Working paper, January 2026, 2026. URL `https://www.jmorenomedina.com/research`.