Dynamic Panel Estimation for Unbalanced and Nonconsecutive Panels: Evidence from the Effect of Crime Perception on Property Prices

Oliver Raab*

August 17, 2026

Abstract

This paper proposes a new estimator for dynamic panel data models with highly sparse, unbalanced panels containing nonconsecutive observations. Existing dynamic panel estimators typically rely either on first-difference transformations, which require consecutive observations, or on analytical bias corrections, which assume balanced or only moderately unbalanced panels. As a result, these methods are generally inapplicable when individual units are observed irregularly over time and contain large gaps between observations. I develop a new dynamic panel estimator that addresses these limitations. The estimator is inspired by the long-difference literature but departs from existing approaches by exploiting the idea that, under certain conditions, individual fixed effects can be eliminated through a causal matching strategy, while the remaining group-specific fixed effects are removed through an additional differencing step. The proposed estimator constructs observation-specific instrument sets based on the nearest available lag used in the differencing step. Consequently, observations with more recent lagged values admit a richer set of valid instruments, whereas observations with larger temporal gaps rely on fewer historical instruments while remaining identified. I illustrate the method by estimating the dynamic effect of crime perception on neighborhood property-price appreciation. The results indicate that increasing crime perception from the 25th to the 75th

*Ph.D. Program in Economics, CUNY Graduate Center; oraab@gradcenter.cuny.edu

percentile is associated with approximately a 0.7% increase in subsequent property-price appreciation. The effect is stronger during the Biden administration than during the Trump administration.

1 Introduction

Adding a lagged dependent variable to a panel-data model is useful when the outcome follows a smooth or gradual adjustment process rather than abruptly moving to a new equilibrium. Dependence on past values is commonly interpreted as persistence. Property prices provide a natural example of this mechanism, because current housing valuations in a neighborhood depend in part on previous valuations. However, including lagged property values as regressors in a panel model with unit fixed effects creates the well-known dynamic-panel bias studied by Nickell (1981). Nickell shows that the least-squares dummy-variable estimator, or equivalently the within-group estimator implied by the Frisch-Waugh-Lovell theorem, is inconsistent due the lagged dependent variable is correlated with the transformed error term in panels with large N and fixed T.
This problem was the catalist for an extensive stream of literature on dynamic panel-data estimation. One strand of this literature removes the problematic fixed effect before estimation through model transformations, including first-difference estimators (Anderson & Hsiao, 1982; Arellano & Bond, 1991) and alternative differencing approaches (Arellano & Bover, 1995; Hahn et al., 2007). Another strand estimates the fixed-effects model directly and then applies analytical bias corrections (Bruno, 2005; Bun & Kiviet, 2003; Kiviet, 1995). These methods are powerful in conventional panel settings, but they are less well suited to applications in which the data contain many missing and nonconsecutive observations. First-difference and long-difference methods require usable lag structures, while many analytical corrections are designed for balanced or only moderately unbalanced panels. These requirements become restrictive when the panel is highly sparse and many units are not observed in consecutive periods.
This paper develops and applies a method for analyzing dynamic panel-data models when observations are sparsely populated over time. The empirical setting is a repeated cross section collapsed into a panel at the spatial unit of interest, in this case ZIP codes. This structure is common in applied econometrics: repeated survey observations are aggregated to local geographic units, but the resulting panel is highly unbalanced because many ZIP codes are not observed in every year and often do not appear in consecutive years. Applying standard first-difference meth-

ods would therefore substantially reduce the usable sample. At the same time, analytical bias-correction methods are difficult to apply because the panel is far from balanced.

The proposed approach is inspired by the long-difference logic in Hahn et al. (2007), but adapts it to the specific structure of the data. The method starts from the idea that the importance of individual fixed effects can be addressed through the causal research design, while remaining group-level fixed effects can be removed through an additional differencing step. I split the sample into treated-period observations and comparison-period observations. For each treated-period observation, I find the most recent available comparison observation within the same bundle-by-DMA cell. If the same ZIP code is available in the comparison period, I use the average of those same-ZIP observations. If no same-ZIP observation is available, I use the average of the most recent comparison observations in the same bundle-by-DMA cell. Differencing the treated observation from this matched comparison observation removes the shared bundle-by-DMA component.

The method also constructs observation-specific instrument sets for the endogenous lagged price change. The available instruments depend on the timing of the matched comparison observation. When the comparison observation is more recent, more historical price levels and price changes are available as instruments; when the comparison observation is further in the past, the instrument set is smaller. This feature allows the estimator to retain observations with irregular time gaps while still using valid historical price information to instrument the lagged dependent variable.

I apply this method to estimate the effect of crime perception on next-year property-price appreciation. Property prices typically follow a gradual adjustment process, making the inclusion of a lagged dependent variable important for isolating the effect of crime perception on future price appreciation. However, the structure of the data presents several challenges. The final dataset, aggregated at the ZIP-code level, contains relatively few observations for many spatial units. To address this issue, I estimate ZIP-code-level crime perception using the Battese–Harter–Fuller small-area estimator, which constructs partially pooled estimates by combining ZIP-specific information with information from the broader sample.

In addition, the resulting panel is highly unbalanced and many ZIP codes are not observed in consecutive years, making conventional dynamic panel estimators difficult to apply. I therefore implement the dynamic panel estimator proposed in this paper. The specification compares Trump-period and Biden-period ZIP-year observations with matched Obama-period observations within the same fixed-effect group. Standard errors are clustered at the Obama-match level and incorporate Murphy–Topel corrections to account for the generated crime-perception regressor.

The results suggest that crime perception is positively associated with future property-price appreciation in the log-price specification, with the effect being larger and more precisely estimated during the Biden administration than during the Trump administration. In the level-price specification, the Trump-period estimate is negative and imprecisely estimated, whereas the Biden-period estimate is positive but statistically insignificant. These findings suggest that ZIP codes with higher perceived crime may subsequently experience stronger property-price appreciation, possibly reflecting neighborhood renewal, increased demand for relatively inexpensive housing, or the process commonly referred to as gentrification. The estimates should therefore be interpreted as evidence from a sparse dynamic-panel setting in which the proposed matching-and-instrumenting strategy preserves substantially more information than conventional estimators based on consecutive-period differencing.

1.1 Literature Review

The inconsistency arising from including a lagged dependent variable in an AR(1) panel model has been studied extensively since Nickell (1981). A large body of literature has emerged proposing methods to address this bias, and this paper contributes to two related strands of that literature.

First, this paper builds on the large literature on transformation-based dynamic panel estimators, including the influential contributions of Anderson and Hsiao (1982), Arellano and Bond (1991), Arellano and Bover (1995), Blundell and Bond (1998), and Hahn et al. (2007). These papers share the common objective of eliminating the problematic individual fixed effects from the estimation procedure while exploiting lagged values of the endogenous variable as instruments. Specifically, Anderson and Hsiao (1982), Arellano and Bond (1991), and Blundell and Bond (1998) rely on first-difference transformations, whereas Arellano and Bover (1995) employs forward orthogonal deviations and Hahn et al. (2007) uses long differences. Each of these approaches also proposes a corresponding set of valid instruments. The estimator proposed in this paper draws on these ideas by combining a long-difference-type transformation with an observation-specific instrument set and a differencing strategy based on group averages, making the estimator applicable to highly sparse panels with nonconsecutive observations.

Second, this paper builds on the analytical bias-correction literature, which extends dynamic panel estimation beyond first-difference-based approaches. In particular, Bruno (2005) extends the LSDV analytical bias correction to unbalanced panels. Nevertheless, the estimator still assumes a standard dynamic panel structure in which each usable observation is accompanied by its immediate lag. This assumption is too restrictive for the sparse, nonconsecutive panels considered in this paper, where many observations are separated by multiple time periods.

The identification strategy also builds on Raab (2026a) and Raab (2026b). These papers provide the spatial bundle structure, the crime-perception measurement framework, and the housing-price setting used in the present application. This paper extends that framework by adding dynamic price adjustment and by instrumenting lagged price appreciation with historical housing-price instruments. The spatial comparison logic is also related to empirical strategies in Kaplan and Yuan (2020), Spenkuch and Toniatti (2018), and Snyder and Strömberg (2010), which use geographically or institutionally induced variation to support causal interpretation.

The empirical application connects this dynamic panel contribution to the literature on crime and housing markets. A long literature studies how crime affects property values (Ihlanfeldt & Mayock, 2010). Many papers find that higher crime is associated with lower property values, including Thaler (1978), Bowes and Ihlanfeldt (2001), and Gibbons (2004). Other studies report positive or statistically insignificant relationships (Buck et al., 1993; Case & Mayer, 1996; Kain & Quigley, 1970; Ridker & Henning, 1967). This paper differs from much of that literature by focusing on crime perception rather than realized crime and by studying property-price appreciation in a sparse dynamic panel.

Finally, the paper relates to work showing that measured crime and perceived crime need not coincide. Official crime statistics may understate crime intensity or vary with reporting behavior (Gibbons, 2004; MacDonald, 2002). Buonanno et al. (2013) study crime perceptions directly in a hedonic property-price model, while Linden and Rockoff (2008) and Pope (2008) examine information shocks from sex-offender registries. The present paper extends this perception-based literature by estimating the dynamic relationship between crime perception and subsequent property-price appreciation in a highly unbalanced ZIP-code panel.

2 The Model

Consider the following autoregressive panel-data model for property-price appreciation:

$$\Delta y_{i,t+1} = \alpha + \rho \, \Delta y_{it} + \beta \, x_{it} + \mathbf{z}'_{it}\boldsymbol{\gamma} + \eta_i + v_{i,t}. \tag{1}$$

Here, x_{it} denotes the regressor of interest, $\Delta y_{i,t+1}$ is the dependent variable measuring property-price appreciation between periods t and $t+1$, Δy_{it} is its lagged value, and $\mathbf{z}_{it}$ is a vector of control variables. The composite error term is given by $\varepsilon_{i,t+1} = \eta_i + v_{i,t+1}$, where η_i denotes the unobserved individual fixed effect and $v_{i,t+1}$ is the idiosyncratic error term. I assume $i = 1, \ldots, N$, $t = 1, \ldots, T$, $|\rho| < 1$, and $v_{i,t+1} \sim \mathcal{N}(0, \sigma_v^2)$.

Collecting observations over time and across individuals gives

$$\mathbf{y} = \mathbf{D}\boldsymbol{\eta} + \mathbf{W}\boldsymbol{\delta} + \boldsymbol{\varepsilon}, \tag{2}$$

where

$$\mathbf{W} = [\Delta\mathbf{y}; \mathbf{X}; \mathbf{Z}],$$

$\mathbf{y}$ and $\mathbf{W}$ are the $(NT \times 1)$ and $(NT \times k)$ matrices of stacked observations, respectively, $\mathbf{D}$ is the $(NT \times N)$ matrix of individual dummy variables, $\boldsymbol{\eta}$ is the $(N \times 1)$ vector of individual fixed effects, $\boldsymbol{\varepsilon}$ is the $(NT \times 1)$ vector of disturbances, and $\boldsymbol{\delta} = [\rho,\ \beta,\ \boldsymbol{\gamma}']'$ is the vector of slope coefficients.

Condition 1. The causal strategy eliminates the importance of individual fixed effects through an appropriate matching strategy. However, while eliminating the nuance of the individual effect, creates a shared group common fixed effect, $\lambda_{g(i)}$, which must be removed for consistency.

Assuming that Condition 1 is met, the model becomes:

$$\Delta y_{i,t+1} = \alpha + \rho \, \Delta y_{it} + \beta \, x_{it} + \mathbf{z}'_{it}\boldsymbol{\gamma} + \lambda_{g(i)} + v_{i,t}. \tag{3}$$

The appropriate causal strategy therefore trades the individual effect for a group fixed effect. However, the strategy that would be sufficient for causal interpretation of β under different circumstances is still not sufficient here, because the group effect creates an endogeneity problem for LSDV similar to the one created by the individual fixed effect.

The existence of a group-specific fixed effect offers a further generalization of the FWL theorem for LSDV with group effects. In other words, Equation (3) is equivalent to a within-group estimator that is run by transforming the same equation (Gaure, 2013). For the proof, refer to Appendix A.

The objective is therefore to eliminate the problematic group fixed effect, similarly to the transformation-based models outlined by Anderson and Hsiao (1982), Arellano and Bond (1991), Arellano and Bover (1995), Blundell and Bond (1998), and Hahn et al. (2007). To do so, I start with a sample split, where one sample is the core sample of interest and the other is an auxiliary sample feeding instruments.

Condition 2. Each observation in the core sample must have a corresponding observation in the auxiliary sample belonging to the same group. In addition, the observation must have at least one earlier lagged observation, preceding the corresponding matched group observation, available for use as an instrument.

After splitting the sample at $t = k$, two subsamples emerge: the auxiliary sample, containing observations from $t = 1, \dots, k$, and the core sample, containing observations from $t = k + 1, \dots, T$. Denote the time periods of the auxiliary and core samples by $t_A \in \{1, \dots, k\}$ and $t_C \in \{k + 1, \dots, T\}$, respectively. Owing to the highly unbalanced nature of the panel, each observation in the core sample is matched to the most recent observation in the auxiliary sample belonging to the same group. Choosing the most recent feasible match maximizes the number of preceding lagged observations available for use as instruments. Therefore, each observation in the core sample is matched to its group-mean counterpart in the auxiliary sample. In other words, observation i from the core sample is matched to the mean of group $g(i)$ at the auxiliary-sample period t_A^* satisfying

$$t_C - t_A^* < t_C - t_A, \qquad \forall\, t_A \neq t_A^*, \quad t_A \in \{1, \dots, k\}.$$

Following the matching step, the remaining group-specific fixed effect is removed by subtracting the corresponding matched group means from all variables. The transformed model is therefore

$$\begin{aligned}\Delta y_{i,t_C+1} - \overline{\Delta y}_{g(i),t_A^*} = \rho\left(\Delta y_{i,t_C} - \overline{\Delta y}_{g(i),t_A^*}\right) + \beta\left(x_{i,t_C} - \bar{x}_{g(i),t_A^*}\right) \\ + \left(\mathbf{z}_{i,t_C} - \bar{\mathbf{z}}_{g(i),t_A^*}\right)' \boldsymbol{\gamma} + \left(v_{i,t_C} - \bar{v}_{g(i),t_A^*}\right).\end{aligned} \tag{4}$$

for

$$t_C = k + 1, \dots, T, \qquad t_A^* \in \{1, \dots, k\}.$$

where

$$\overline{\Delta y}_{g(i),t_A^*} = \frac{1}{n_{g(i),t_A^*}} \sum_{j \in g(i)} \Delta y_{j,t_A^*}, \tag{5}$$

$$\bar{x}_{g(i),t_A^*} = \frac{1}{n_{g(i),t_A^*}} \sum_{j \in g(i)} x_{j,t_A^*}, \tag{6}$$

$$\bar{\mathbf{z}}_{g(i),t_A^*} = \frac{1}{n_{g(i),t_A^*}} \sum_{j \in g(i)} \mathbf{z}_{j,t_A^*}, \tag{7}$$

$$\bar{v}_{g(i),t_A^*} = \frac{1}{n_{g(i),t_A^*}} \sum_{j \in g(i)} v_{j,t_A^*}. \tag{8}$$

Here, $n_{g(i),t_A^*}$ denotes the number of observations belonging to group $g(i)$ at the matched auxiliary-sample period t_A^*.

Matching to the closest group mean from the auxiliary sample can potentially offer a rich array of instruments, depending on the length of the auxiliary sample. The instruments are based on the orthogonality condition

$$E\left[\Delta y_{i,t_A}\left(v_{i,t_C} - \bar{v}_{g(i),t_A^*}\right)\right] = 0$$

for all $t_A \neq t_A^*$ and $t_A < t_A^*$. Therefore, depending on the matched period t_A^* for each observation in the core sample, there is a different number of potential instruments for the endogenous lagged regressor Δy_{it}.

3 Monte Carlo experiment

4 Empirical Strategy

For the details on data contruction and U.S. media markets institutional overview, reffer to Raab (2026a) and Raab (2026b). I will briefly outline the identification strategy based on the mentioned papers.

4.1 Identification Strategy

The identification strategy exploits spatial discontinuities at DMA borders by comparing geographically proximate observations assigned to different media markets. Households located near opposite sides of the same DMA boundary are expected to share similar local economic conditions, housing-market trends, amenities, and demographic characteristics while being exposed to different local television programming. Restricting the analysis to neighborhoods near the same DMA boundary and controlling for spatial-bundle fixed effects therefore isolates plausibly exogenous variation in media exposure arising from DMA assignment rather than from systematic differences across locations.

4.2 Empirical Framework

In this section, I illustrate the proposed estimator by estimating the effect of crime perception on future property-price appreciation while controlling for current appreciation to account for price persistence. Because crime perception is measured using survey responses, many ZIP-year cells contain only a small number of observations, making direct ZIP-level estimates noisy and unreliable. To address this issue, I first estimate neighborhood crime perception using the Battese–Harter–Fuller (BHF) small-area estimator, which combines ZIP-specific information with information from the full sample through partial pooling to obtain more precise ZIP-year estimates. Since the resulting crime-perception measure is a generated regressor obtained in a first-stage estimation, conventional standard errors from the second-stage regression would understate estimation uncertainty. I therefore apply Murphy–Topel standard-error corrections to account for the additional sampling variability introduced by the first-stage estimation. I then estimate the following dynamic model:

$$\Delta \ln P_{it+1} = \alpha + \rho \Delta \ln P_{it} + \beta \widetilde{CP}_{it} + \mathbf{z}_{it}' \boldsymbol{\gamma} + \delta_b + \varepsilon_{it}. \tag{9}$$

where $\Delta \ln P_{it+1}$ denotes appreciation in neighborhood property prices between periods t and $t+1$, $\Delta \ln P_{it}$ is the corresponding lagged appreciation capturing price persistence, $\widetilde{CP}_{it}$ is the Battese–Harter–Fuller estimate of neighborhood crime perception, $\mathbf{z}_{it}$ is a vector of control variables, δ_b denotes spatial-bundle fixed effects, and ε_{it} is the error term. Because crime perception is estimated in a first-stage procedure, I apply Murphy–Topel standard-error corrections to account for the generated-regressor problem. The dynamic specification is estimated using the proposed matching-and-instrumenting estimator described in the previous section. My sample contains data spanning three presidential administrations: Obama, Trump, and Biden. Because the Obama administration covers a longer period, I use the Obama period as the auxiliary sample for two core samples: the Trump administration and the Biden administration. This choice maximizes the number of potential matches across groups. I therefore run two separate models:

$$\begin{aligned}\Delta \ln P^{(a)}_{i,t+1} &= \alpha^{(a)} + \rho^{(a)} \Delta \ln P_{it} + \beta^{(a)} \widetilde{CP}_{it} + \mathbf{z}'_{it}\boldsymbol{\gamma}^{(a)} \\ &\quad + \delta^{(a)}_b + \varepsilon^{(a)}_{it}, \qquad a \in \{\text{Trump}, \text{Biden}\}.\end{aligned} \tag{10}$$

I match each observation in both subsamples, $a \in \{\text{Trump}, \text{Biden}\}$, to its group mean from the Obama-administration auxiliary sample, under the condition that the match is the most recent available observation during the Obama period.
In addition, I include one additional condition stemming from the specifics of the sample. Because I match at the ZIP-code level, I add a preference for direct ZIP-code matches. In other words, if a ZIP-code counterpart is available in the Obama sample, I prefer it as the match. If a ZIP-code match is not available, then I match the ZIP code to the group average, as outlined in the previous section.
The advantage of the sample is that property-price appreciation is available for each period and is therefore balanced. Crime perception, however, is not available for each ZIP-year, which makes the estimation sample sparsely populated and unbalanced. The availability of property prices offers a rich set of instruments for observations available in each t_C. For example, if t^*_A is the matched year during the Obama period, then each prior year can be used as an instrument. Therefore, there are $t^*_A - t_1$ available instruments in levels.

5 Results

Table 1 reports the dynamic panel estimates of the relationship between crime perception and future property-price appreciation. Focusing on the log-price specifications, higher crime perception is positively associated with subsequent property-price appreciation. Moving crime perception from the 25th to the 75th percentile is associated with approximately a 0.5 percent increase in future appreciation during the Trump administration and a 0.7 percent increase during the Biden administration. The Biden-period estimate is larger and much more significant. The log specifications also show substantial persistence in property-price appreciation, with positive and statistically significant coefficients on lagged price appreciation in both periods.

Table 1: Dynamic Panel Specification: Crime Perception and Property Price Appreciation

	Thousands of $		Logs	
	Trump	Biden	Trump	Biden
Dependent Variable : Future Property Price Appreciation				
Intercept	0.111***	0.085***	0.024***	0.006
	(0.042)	(0.032)	(0.004)	(0.004)
Crime Perception	-0.016	0.011	0.005*	0.007***
	(0.013)	(0.012)	(0.003)	(0.003)
Δ Lagged Price	-0.510	0.530***	0.470***	0.548***
	(1.110)	(0.171)	(0.114)	(0.074)
N	964	572	964	572
Unique Treated ZIPs	855	536	855	536
Obama Match Clusters	364	274	364	274
First-stage Statistic	93.312	216.799	78.827	89.848
Controls	Yes	Yes	Yes	Yes
Obama Match Cluster SE	Yes	Yes	Yes	Yes
Murphy-Topel Correction	Yes	Yes	Yes	Yes

Notes: Each column compares treated-period ZIP-years to matched Obama-period ZIP-years in the same bundle-by-DMA cell. The Obama comparison is the most recent available Obama year. Same-ZIP matches are averaged when available; otherwise, same bundle-by-DMA Obama observations are averaged. The dependent variable is the treated future price change minus the matched Obama future price change. The first two columns measure price changes in hundreds of thousands of dollars; the last two columns measure log price changes. The endogenous regressor is the corresponding difference in lagged price changes. It is instrumented with imputed historical Zillow price levels and price changes from 2007 onward. Missing ZIP prices are filled by adjacent ZIP averages, county prices, and then state prices. Crime perception is the BHF/EBLUP prediction, centered at the 25th percentile and scaled by the interquartile range before matching. Controls are differenced mean-centered ACS DP04 property controls and are suppressed from the table. Standard errors are clustered by the Obama match cluster and include the Murphy-Topel generated-regressor correction. The first-stage statistic is the linearmodels first-stage Wald statistic for the excluded instruments. * p<0.10, ** p<0.05, *** p<0.01.

6 Conclusion

A FWL Proof

Consider the model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}_g\boldsymbol{\lambda} + \mathbf{u}, \tag{11}$$

where $\mathbf{D}_g \in \mathbb{R}^{n\times G}$ is the matrix of group dummy variables.

Define the projection matrix

$$\mathbf{P}_g = \mathbf{D}_g(\mathbf{D}_g'\mathbf{D}_g)^{-1}\mathbf{D}_g'$$

and the residual-maker matrix

$$\mathbf{M}_g = \mathbf{I} - \mathbf{P}_g = \mathbf{I} - \mathbf{D}_g(\mathbf{D}_g'\mathbf{D}_g)^{-1}\mathbf{D}_g'.$$

The Frisch–Waugh–Lovell theorem therefore implies

$$\hat{\boldsymbol{\beta}}_{\text{LSDV}} = (\mathbf{X}'\mathbf{M}_g\mathbf{X})^{-1}\mathbf{X}'\mathbf{M}_g\mathbf{y}.$$

Consider,

$$\mathbf{M}_g\mathbf{y} = \mathbf{y} - \mathbf{P}_g\mathbf{y},$$

whose ith element is

$$(\mathbf{M}_g\mathbf{y})_i = y_i - \bar{y}_{g(i)}.$$

Similarly:

$$(\mathbf{M}_g\mathbf{X})_i = x_i - \bar{x}_{g(i)}.$$

Where $\bar{y}_{g(i)}$ and $\bar{x}_{g(i)}$ are computed across all individuals and time periods within group $g(i)$. Premultiplying the original model by $\mathbf{M}_g$ yields

$$\mathbf{M}_g\mathbf{y} = \mathbf{M}_g\mathbf{X}\boldsymbol{\beta} + \mathbf{M}_g\mathbf{D}_g\boldsymbol{\lambda} + \mathbf{M}_g\mathbf{u}.$$

the transformed model becomes

$$\mathbf{M}_g\mathbf{y} = \mathbf{M}_g\mathbf{X}\boldsymbol{\beta} + \mathbf{M}_g\mathbf{u}.$$

Where,

$$y_i - \bar{y}_{g(i)} = (x_i - \bar{x}_{g(i)})'\boldsymbol{\beta} + (u_i - \bar{u}_{g(i)}).$$

and

$$\hat{\boldsymbol{\beta}}_{\text{group-within}} = (\mathbf{X}'\mathbf{M}_g\mathbf{X})^{-1}\mathbf{X}'\mathbf{M}_g\mathbf{y}.$$

Therefore,

$$\hat{\boldsymbol{\beta}}_{\text{LSDV}} = \hat{\boldsymbol{\beta}}_{\text{group-within}}.$$

B Additional Regressions

Table 2: Effect of Crime Perception on Neighborhood Property Price Appreciation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable : First Difference in Neighborhood Property Prices (hundreds of thousands of dollars)								
Intercept	0.147***	0.070***	0.149**	0.018	0.090***	0.005	-0.033	-0.041
	(0.028)	(0.017)	(0.060)	(0.052)	(0.020)	(0.026)	(0.076)	(0.073)
Trump		0.204***		0.243***		0.242***		0.255**
		(0.062)		(0.065)		(0.062)		(0.099)
Biden		0.063**		0.094***		0.086***		0.144
		(0.027)		(0.027)		(0.033)		(0.091)
Crime Perception	-0.084***	-0.020	-0.039***	0.044*	-0.025**	0.051**	-0.037***	0.041*
	(0.027)	(0.019)	(0.014)	(0.023)	(0.011)	(0.024)	(0.012)	(0.022)
Crime Perception × Trump		-0.171**		-0.224***		-0.221***		-0.204***
		(0.080)		(0.081)		(0.076)		(0.076)
Crime Perception × Biden		-0.033		-0.054**		-0.038		-0.058**
		(0.032)		(0.026)		(0.035)		(0.029)
Lagged Price Change	0.406***	0.414***	0.308***	0.320***	0.369***	0.374***	0.275***	0.286***
	(0.064)	(0.069)	(0.060)	(0.054)	(0.055)	(0.055)	(0.065)	(0.063)
N	3,050	3,050	3,050	3,050	3,050	3,050	3,050	3,050
Bundles	153	153	153	153	153	153	153	153
DMAs	160	160	160	160	160	160	160	160
Years	14	14	14	14	14	14	14	14
Gallup Observations	3,174	3,174	3,174	3,174	3,174	3,174	3,174	3,174
Bundle Fixed Effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
DMA Fixed Effects	No	No	No	No	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	No	No	No	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weights	Distance	Distance	Distance	Distance	Distance	Distance	Distance	Distance
Presidential Interactions	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Each observation is a ZIP-year area. The dependent variable is a first difference in future property prices, defined as $P_{t+1} - P_t$. The lagged price-change control is $P_t - P_{t-1}$. Crime perception is a BHF/EBLUP prediction from the unit-level Gallup model, centered at the 25th percentile and scaled by the interquartile range. Second-stage standard errors apply a Murphy-Topel correction for the generated crime-perception regressor. Controls include mean-centered ACS DP04 bedroom shares, year-built shares, units in structure, and vacancy variables. Omitted categories are 3 bedrooms, built 1980-1999, and mobile homes; median rooms are excluded. Missing Census controls are filled from the nearest same-year ZCTA with complete controls. The base second-stage covariance is HC1 robust; ZIP-year weights use only the DMA-border distance weight. Columns (7) and (8) add year fixed effects. * p<0.10, ** p<0.05, *** p<0.01.

Table 3: Effect of Crime Perception on Neighborhood Property Price Appreciation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable : First Difference in Log Neighborhood Property Prices								
Intercept	0.026***	0.015***	0.070*	0.053	0.024***	0.015**	-0.019*	-0.019*
	(0.003)	(0.005)	(0.042)	(0.040)	(0.008)	(0.008)	(0.011)	(0.011)
Trump		0.036***		0.035***		0.032***		0.065***
		(0.008)		(0.006)		(0.005)		(0.011)
Biden		0.002		0.004		-0.001		0.043***
		(0.007)		(0.006)		(0.005)		(0.010)
Crime Perception	0.001	0.003	0.007**	0.010**	0.008***	0.010**	0.006**	0.006*
	(0.003)	(0.005)	(0.003)	(0.004)	(0.003)	(0.004)	(0.002)	(0.003)
Crime Perception × Trump		-0.013**		-0.014***		-0.012**		-0.000
		(0.006)		(0.005)		(0.005)		(0.004)
Crime Perception × Biden		0.001		-0.000		0.001		-0.000
		(0.005)		(0.005)		(0.005)		(0.005)
Lagged Log Price Change	0.387***	0.383***	0.369***	0.368***	0.393***	0.391***	0.210***	0.210***
	(0.049)	(0.049)	(0.045)	(0.045)	(0.042)	(0.042)	(0.038)	(0.038)
N	3,050	3,050	3,050	3,050	3,050	3,050	3,050	3,050
Bundles	153	153	153	153	153	153	153	153
DMAs	160	160	160	160	160	160	160	160
Years	14	14	14	14	14	14	14	14
Gallup Observations	3,174	3,174	3,174	3,174	3,174	3,174	3,174	3,174
Bundle Fixed Effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
DMA Fixed Effects	No	No	No	No	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	No	No	No	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weights	Distance	Distance	Distance	Distance	Distance	Distance	Distance	Distance
Presidential Interactions	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Each observation is a ZIP-year area. The dependent variable is a first difference in future property prices, defined as $P_{t+1} - P_t$. The lagged price-change control is $P_t - P_{t-1}$. Crime perception is a BHF/EBLUP prediction from the unit-level Gallup model, centered at the 25th percentile and scaled by the interquartile range. Second-stage standard errors apply a Murphy-Topel correction for the generated crime-perception regressor. Controls include mean-centered ACS DP04 bedroom shares, year-built shares, units in structure, and vacancy variables. Omitted categories are 3 bedrooms, built 1980-1999, and mobile homes; median rooms are excluded. Missing Census controls are filled from the nearest same-year ZCTA with complete controls. The base second-stage covariance is HC1 robust; ZIP-year weights use only the DMA-border distance weight. Columns (7) and (8) add year fixed effects. * p<0.10, ** p<0.05, *** p<0.01.