

# Building Demographics and Eviction Intensity

Philly Evictions Project

2026-02-20

## Table of contents

<b>1 Overview</b>	<b>1</b>
<b>2 Eviction Intensity Classification</b>	<b>2</b>
<b>3 Descriptive Means</b>	<b>2</b>
<b>4 Regression Specifications</b>	<b>2</b>
4.1 (A) Unconditional . . . . .	2
4.2 (B) Tract FE + Controls . . . . .	3
4.3 (C) Block Group FE + Controls . . . . .	3
4.4 (D) Hurdle Spline . . . . .	3
<b>5 Results</b>	<b>3</b>
5.1 Main Table: All Outcomes, Tract and BG FE . . . . .	3
5.2 Full Specifications: Share Black . . . . .	4
5.3 Coefficient Plots . . . . .	4
<b>6 Interpretation</b>	<b>4</b>

## 1 Overview

This document summarizes the building demographics analysis from `r/building-demographics-evictions.R`. The script asks: **are high-evicting buildings disproportionately occupied by Black tenants, female-headed households, or Black female-headed households?**

**Unit of analysis:** Building (PID), collapsed to a single cross-section by taking unit-weighted means of demographic shares across all available years.

### Key data inputs:

- `bldg_panel_blp` — building  $\times$  year panel with tenant demographics (from InfoUSA imputations), eviction filing rates, and building characteristics

### Demographic outcomes:

- `infousa_pct_black` — share of tenants imputed as Black
- `infousa_pct_female` — share of head-of-household imputed as female
- `infousa_pct_black_female` — share of tenants who are both Black and female HoH

### Sample restrictions:

- Demographic coverage  $\geq 0.5$  (`infousa_share_persons_demog_ok`)
- Non-missing GEOID, total area  $> 0$ , market value  $> 0$
- EB filing rate  $\leq 0.75$  (excludes extreme outliers)

## 2 Eviction Intensity Classification

Buildings are classified using the empirical Bayes eviction filing rate (pre-COVID), with true never-filers separated:

Bin	Definition
No filings	<code>total_filings_pre2019 == 0</code> (true zero — distinct from low EB rate)
(0–5%]	EB rate $\leq 0.05$ , at least 1 filing pre-2019
(5–10%]	EB rate $\in (0.05, 0.10]$ (reference category)
(10–20%]	EB rate $\in (0.10, 0.20]$
20%+	EB rate $> 0.20$

**Note:** The EB filing rate is a posterior mean from an empirical Bayes procedure, so it is always  $> 0$  even for buildings with zero observed filings. We use the raw filing count (`total_filings_pre2019`) to identify true never-filers.

## 3 Descriptive Means

Table 2: Unit-weighted mean tenant demographics by eviction intensity bin

Eviction bin	N buildings	Mean units	% Black	% Female	% Black Female	Demog coverage
No filings	101,190	9.7	31.3	55.1	19.0	0.936
(0–5%]	4,688	89.9	23.0	57.2	15.1	0.919
(5–10%]	20,834	23.2	42.1	57.7	26.4	0.921
(10–20%]	13,506	26.4	45.9	60.1	29.4	0.918
20%+	5,385	60.4	48.9	61.4	32.1	0.907

## 4 Regression Specifications

For each demographic outcome  $Y_i$ , four specifications are estimated:

### 4.1 (A) Unconditional

$$Y_i = \sum_b \beta_b \cdot \mathbb{1}[\text{bin}_i = b] + \varepsilon_i$$

Unit-weighted, clustered by census tract.

## 4.2 (B) Tract FE + Controls

$$Y_i = \sum_b \beta_b \cdot \mathbb{1}[\text{bin}_i = b] + \gamma \log(\text{area}_i) + \delta \log(\text{value}_i) + \alpha_{\text{tract}} + \phi_X + \varepsilon_i$$

where  $\phi_X$  absorbs FE for unit-count bin, building type, construction decade, quality grade, and number-of-stories bin. Clustered by tract.

## 4.3 (C) Block Group FE + Controls

Same as (B) but with block group FE instead of tract FE. Clustered by GEOID.

## 4.4 (D) Hurdle Spline

Replaces the eviction bins with a flexible functional form:

$$Y_i = \beta_0 \cdot \text{NeverFiled}_i + \sum_{s=1}^6 \beta_s \cdot B_s(\text{EB rate}_i) + \text{controls} + \alpha_{\text{BG}} + \phi_X + \varepsilon_i$$

where  $B_s(\cdot)$  are natural spline basis functions (df = 6) of the EB filing rate for ever-filers, and the spline basis is set to zero for never-filers (the “hurdle”). This allows the relationship between eviction intensity and demographics to be nonlinear without imposing bin boundaries.

# 5 Results

## 5.1 Main Table: All Outcomes, Tract and BG FE

Dependent Variables:	Share Black		Share Female		Share Black Female	
Model:	Tract FE	BG FE	Tract FE	BG FE	Tract FE	BG FE
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Variables</i>						
Eviction intensity (EB, pre-COVID) = Nofilings5-10%])	-0.0189*** (0.0025)	-0.0174*** (0.0018)	-0.0120*** (0.0028)	-0.0125*** (0.0026)	-0.0173*** (0.0022)	-0.0167*** (0.0018)
Eviction intensity (EB, pre-COVID) = (0-5%]5-10%])	-0.0066 (0.0051)	-0.0065* (0.0037)	0.0122** (0.0061)	0.0115* (0.0062)	0.0012 (0.0047)	0.0014 (0.0037)
Eviction intensity (EB, pre-COVID) = (10-20%]5-10%])	0.0205*** (0.0043)	0.0149*** (0.0028)	0.0151** (0.0064)	0.0134*** (0.0043)	0.0182*** (0.0035)	0.0153*** (0.0027)
Eviction intensity (EB, pre-COVID) = 20%+5-10%])	0.0291*** (0.0074)	0.0330*** (0.0051)	0.0204*** (0.0059)	0.0244*** (0.0058)	0.0301*** (0.0057)	0.0340*** (0.0046)
<i>Fixed-effects</i>						
census_tract	Yes		Yes		Yes	
num_units_bin	Yes	Yes	Yes	Yes	Yes	Yes
building_type	Yes	Yes	Yes	Yes	Yes	Yes
year_blt_decade	Yes	Yes	Yes	Yes	Yes	Yes
quality_grade_standard	Yes	Yes	Yes	Yes	Yes	Yes
num_stories_bin	Yes	Yes	Yes	Yes	Yes	Yes
GEOID		Yes		Yes		Yes
<i>Fit statistics</i>						
Observations	145,599	145,593	145,599	145,593	145,599	145,593
R <sup>2</sup>	0.80906	0.88069	0.06114	0.07723	0.61581	0.66879

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

## 5.2 Full Specifications: Share Black

Dependent Variable:	Share Black			
Model:	Unconditional (1)	Tract FE (2)	BG FE (3)	Spline (4)
<i>Variables</i>				
Eviction intensity (EB, pre-COVID) = No filings 5-10%]	-0.1079*** (0.0133)	-0.0189*** (0.0025)	-0.0174*** (0.0018)	
Eviction intensity (EB, pre-COVID) = (0-5%] 5-10%]	-0.1911*** (0.0236)	-0.0066 (0.0051)	-0.0065* (0.0037)	
Eviction intensity (EB, pre-COVID) = (10-20%] 5-10%]	0.0382** (0.0169)	0.0205*** (0.0043)	0.0149*** (0.0028)	
Eviction intensity (EB, pre-COVID) = 20%+ 5-10%]	0.0683*** (0.0244)	0.0291*** (0.0074)	0.0330*** (0.0051)	
<i>Fixed-effects</i>				
census_tract		Yes		
num_units_bin		Yes	Yes	Yes
building_type		Yes	Yes	Yes
year_blt_decade		Yes	Yes	Yes
quality_grade_standard		Yes	Yes	Yes
num_stories_bin		Yes	Yes	Yes
GEOID			Yes	Yes
<i>Fit statistics</i>				
Observations	145,603	145,599	145,593	145,593
R <sup>2</sup>	0.03605	0.80906	0.88069	0.88072

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

## 5.3 Coefficient Plots

Coefficient plot not found.

All-outcomes coefficient plot not found.

## 6 Interpretation

The analysis tests whether high-evicting buildings have a systematically different tenant composition than comparable buildings in the same neighborhood. By conditioning on block group (or tract) and building observables, the regressions isolate *within-neighborhood* variation in eviction intensity.

Key patterns to look for:

- **Monotonic gradient:** Do the coefficients increase monotonically with eviction intensity? This would indicate that higher-evicting buildings systematically house more Black/female tenants even within the same neighborhood.
- **Never-filer anomaly:** The “No filings” bin captures buildings with zero pre-2019 filings. If these buildings look different from the low-rate (0–5%] bin, it may reflect a structural difference (e.g., owner-occupied buildings misclassified as rentals, or buildings too small to generate filings).
- **Spline nonlinearity:** The hurdle spline (spec D) allows the relationship to curve. If the

effect is concentrated at the top of the distribution, the spline coefficients will be large only for high basis values.