Wildfire Smoke and Voting Behavior: Evidence from California Census Tracts

Working Draft

February 15, 2026

Abstract

This report extends the national county-level analysis of wildfire smoke effects on voting behavior to census tract-level resolution within California. Using Stanford ECHO Lab daily tract-level smoke $PM_{2.5}$ predictions (2006–2022) and precinct-level election returns allocated to census tracts, we estimate two-way fixed effects (TWFE) panel regressions at approximately 8,000 CA tract units. The tract-level analysis exploits finer within-county variation in smoke exposure, enabling identification from spatial variation that is absorbed at the county level. We compare estimates with the national county-level results and examine robustness to county×year fixed effects, alternative smoke thresholds, and crosswalk methodology.

Contents

1   Introduction

The national analysis establishes relationships between pre-election wildfire smoke exposure and voting outcomes at the county level using a TWFE design. This companion analysis pursues three goals:

1. **Finer spatial resolution**: Census tracts ($\sim$4,000 residents) offer much finer variation in both smoke exposure and voting outcomes than counties ($\sim$100,000 residents). Within-county variation in smoke exposure—driven by topography, wind patterns, and proximity to fire sources—provides identifying variation that county-level data cannot capture.

2. **Within-California variation**: Restricting to a single state eliminates state-level confounders entirely, as state policy, campaign spending, and voter registration rules are constant across all tracts. County trends and county×year fixed effects serve as the demanding robustness checks (analogous to state trends and state×year FE in the national analysis).

3. **Direct comparison**: Matching the national analysis structure—build-up table, temporal dynamics, threshold comparisons, House elections—enables direct coefficient comparison and assessment of scale-dependence in the estimated effects.

California is an ideal setting for this analysis: it experiences the most wildfire smoke in the continental US, has substantial within-state variation in exposure (from the heavily smoked northern and central regions to the relatively clear coastal and southern tracts), and maintains high-quality precinct-level election data through the Statewide Database.

2   Data

2.1   Smoke Exposure

We use the Stanford ECHO Lab v2 daily tract-level smoke $PM_{2.5}$ predictions (**?**), which provide estimates of wildfire-attributable $PM_{2.5}$ for every census tract in the United States from 2006 through 2023. We filter to California tracts (GEOID prefix 06), yielding approximately 8,000 tracts with daily observations.

For each election, we compute smoke exposure in cumulative pre-election windows of 7, 14, 21, 28, 30, 60, 90 days and the full fire season (June 1 through Election Day). Within each window we calculate:

- Mean smoke $PM_{2.5}$ (primary treatment variable)

- Number of smoke days (any positive smoke $PM_{2.5}$)

- Fraction of days exceeding the haze threshold (>20 $\mu g/m^3$)

- Fraction exceeding "Unhealthy for Sensitive Groups" (>35.5 $\mu g/m^3$)

- Fraction exceeding "Unhealthy" (>55.5 $\mu g/m^3$)

2.2 Election Data

Precinct-level election returns are obtained from the California Statewide Database (SWDB) for general elections 2006–2022. We include:

- Presidential races: 2008, 2012, 2016, 2020

- US House races: 2006–2022 (biennial)

Precinct votes are allocated to census tracts using a pre-built precinct-to-tract crosswalk from the 2025 *Nature Scientific Data* paper, supplemented by areal interpolation from SWDB precinct shapefiles for validation.

California's top-two primary system (effective 2012) means some House races feature two candidates of the same party. We flag these races and exclude them from Democratic vote share analysis (turnout remains valid).

2.3 Controls

Tract-level controls from the American Community Survey (ACS) 5-year estimates:

- Unemployment rate (B23025)

- Log median household income (B19013)

- Log total population (B01003)

- Educational attainment: percent bachelor's degree or higher (B15003)

- Race/ethnicity composition (B03002)

Weather controls from PRISM 4km monthly rasters:

- October mean temperature

- October total precipitation

2.4 Crosswalk Methodology

The precinct-to-tract allocation uses a two-step approach:

1. **Primary**: Pre-built allocation factors from published crosswalk datasets provide statistically validated weights for distributing precinct votes across census tracts.

2. **Validation**: Independent area-weighted overlay of SWDB precinct boundary shapefiles with Census TIGER tract polygons, using population-weighted areal interpolation where possible. Correlation between pre-built and SWDB-derived tract totals should exceed 0.95.

Tract boundary changes across Census decades (2000 vs. 2010 vs. 2020 tracts) are handled by using 2010 Census tract definitions as the primary geography and applying NHGIS crosswalks where needed.

3 Empirical Strategy

Our baseline specification is a TWFE panel regression:

$$Y_{it} = \beta \cdot \text{Smoke}_{it} + \mathbf{X}'_{it}\gamma + \alpha_i + \delta_t + \varepsilon_{it} \tag{1}$$

where Y_{it} is the outcome for tract i in election t, Smoke_{it} is the pre-election smoke exposure measure, $\mathbf{X}_{it}$ is a vector of time-varying controls, α_i are tract fixed effects, and δ_t are year fixed effects. Standard errors are clustered by tract.

Outcomes:

- Democratic two-party vote share

- Incumbent party vote share

- Log total votes (turnout)

Build-up specifications:

1. Raw OLS (no fixed effects)

2. Tract + Year FE (TWFE baseline)

3. TWFE + Controls (main specification)

4. TWFE + Controls + County trends

Identifying variation: With tract and year FE, identification comes from within-tract, across-election variation in smoke exposure. This isolates the effect of differential pre-election smoke from permanent tract characteristics and common time shocks.

Robustness checks:

- County×year FE: most demanding, absorbs all county-level time-varying confounders

- Alternative clustering: by county (accounts for within-county spatial correlation)

- Threshold comparisons: haze vs. USG vs. unhealthy

- Crosswalk sensitivity: re-estimate with independently constructed SWDB crosswalk

- Controls robustness: sequential addition/removal of individual controls

4 Results

4.1 Build-Up Table

[Table 1: Build-up specification table — to be populated from analysis output]

The build-up table presents presidential election results (30-day window) across four specifications. At the tract level, the TWFE specification absorbs permanent tract-level confounders (e.g., baseline partisanship, demographics, geography), leaving within-tract variation over time for identification.

4.2 Temporal Dynamics

[Figure 1: Cumulative temporal dynamics with controls — to be inserted from ca_temporal_7day_mean_cont

The temporal dynamics figures show how estimated effects evolve as the smoke exposure window expands from 7 to 91 days before the election, matching the national analysis format.

4.3 Smoke Exposure Maps

[Figure 2: CA tract-level smoke exposure by year — to be inserted from ca_smoke_map_by_year.png]
[Figure 3: Residualized smoke map — to be inserted from ca_smoke_map_residualized.png]

5 Comparison with National County-Level Results

[Table 2: Side-by-side coefficients, national county vs. CA tract — to be populated]

[Figure 4: Coefficient comparison plot — to be inserted from ca_national_comparison.png]

The comparison table presents Specification 3 (TWFE + controls, 30-day window) estimates from both the national county-level and CA tract-level analyses. Key questions:

- Are the estimates directionally consistent?

- Are tract-level estimates larger or smaller in magnitude? (Finer resolution may attenuate if effects are spatially correlated, or amplify if county averaging dilutes localized effects.)

- Do the relative magnitudes across outcomes (DEM share vs. incumbent vs. turnout) maintain the same pattern?

6   House Elections

[Table 3: House vs. presidential comparison — to be populated]

The House analysis parallels the presidential analysis but covers biennial elections 2006–2022, providing more temporal variation. Top-two primary races (post-2012) where both candidates share a party are excluded from vote share analysis but retained for turnout.

7   Robustness

7.1   County×Year Fixed Effects

County×year FE absorb all county-level time-varying confounders—including county-specific campaign activities, local economic shocks, and county-level smoke patterns. Identification relies solely on within-county, within-year variation across tracts. This is the most demanding specification and analogous to state×year FE in the national analysis.

7.2   Threshold Comparison

[Table 4: Threshold comparison (haze / USG / unhealthy) — to be populated]

7.3   Controls Robustness

[Table 5: Leave-one-out controls sensitivity — to be populated]

7.4 Crosswalk Validation

We compare results using the pre-built crosswalk against results from the independently constructed SWDB areal interpolation crosswalk. If coefficients are robust to the crosswalk method, this mitigates concerns about measurement error in the precinct-to-tract allocation.

7.5 Drop-2020 Analysis

[Figure 5: Full sample vs. excluding 2020 temporal dynamics — to be inserted from ca_temporal_drop2020.p

8 Discussion

The tract-level analysis for California provides several advantages over the national county-level analysis:

1. **Finer variation**: Exploits within-county spatial variation in smoke exposure driven by topography and wind patterns.

2. **State controls**: Eliminates state-level confounders entirely.

3. **More demanding FE**: County×year FE is feasible and highly demanding.

4. **Richer controls**: Tract-level ACS provides granular demographic information.

Limitations include:

1. **Single state**: Results may not generalize beyond California.

2. **Crosswalk uncertainty**: Precinct-to-tract allocation introduces measurement error.

3. **Tract boundary changes**: Comparisons across Census decades involve some geographic imprecision.

4. **Top-two primary**: Post-2012 same-party House races limit House vote share analysis.