Prosecutorial Policy Analysis

AI-Powered Measurement of Criminal Justice Reform in California

Dvir Yogev

February 2026

Center for Law & Justice, UC Berkeley School of Law

Thank You for the Second Look

Last time, you raised two concerns: whether **outcome data** could support the causal questions, and whether the project was **targeted enough** toward specific causal research questions.

Those concerns were fair. Here is what has changed:

1. **We linked outcome data and found a real signal** — pilot results survive year fixed effects; pretrial detention shows a clean causal signal ($p_{\text{pre-trend}} = 0.90$)
2. **We identified 6 specific causal designs** aligned to specific research questions — each one mapped to the exact outcome data source it requires
3. **We restructured the ask** so that each funded phase directly unlocks a specific causal study, not just more infrastructure

The next 15 minutes will walk you through the evidence.

The Measurement Gap

District Attorneys are among the **most powerful actors** in the criminal justice system, yet we lack systematic measurement of:

- How their policies **vary across jurisdictions**
- How policies **change over time**
- Whether stated policy intent **affects outcomes**

Existing research relies on case outcomes or campaign rhetoric—neither captures the *stated policy intent* that guides line prosecutors daily.

What's Missing

No systematic, comparable measurement of internal DA policy documents exists **anywhere**.

What We Built

The first large-scale, AI-coded dataset of internal DA policy documents—infrastructure for causal inference about prosecutorial ideology.

Data: ACLU NorCal Public Records Archive

2,665 internal DA policy documents obtained by ACLU of Northern California via the California Racial Justice Act.

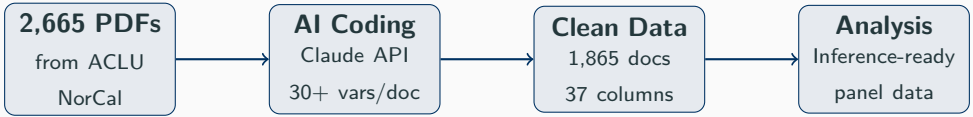
- **41 of 58** California counties represented
- Charging directives, sentencing memos, diversion protocols, racial equity initiatives, bail reform orders
- Documents span 2015–2024

These are *internal operating documents*—not press releases or campaign materials. They reveal what prosecutors actually tell their staff to do.

Why This Is Rare

Internal policy memos are almost never publicly available. The ACLU's PRA effort created a **unique research opportunity** that may not persist indefinitely.

What We Built: Research Pipeline



Dimension Coded	What's Measured
Ideological Orientation	7-point scale: clearly progressive → clearly traditional
Extensive Margin	Impact on *who enters* the system (charging, diversion)
Intensive Margin	Impact on *how severely* people are treated (sentencing)
Specific Policies	Diversion, bail reform, enhancements, three strikes, racial justice
Administrative Context	New policy vs. continuation, DA administration

Total API cost: ~$80 · Processing time: ~2 hours

Not Just "Progressive/Traditional" — We Code Specific Policies

Dimension	Values	Example from Our Data
Bail position	reform / high bail	Placer: reform-oriented bail memo; SLO: bail guideline update
Diversion	yes / no	Sacramento, Santa Clara: explicit diversion program support
Enhancements	minimize / maximize	SF, LA: minimize enhancements; 6 counties maximize
Three strikes	restrictive / expansive	Tracked per-document across all counties
Racial justice	high / moderate / low	9 counties high (post-2020 surge), 16 moderate, 10 low
Juvenile transfer	restrictive / expansive	4 documents with explicit positions
Alt. to incarceration	yes / no	Whether doc endorses treatment, community service, etc.

This is v1.0. Adding new dimensions (gun policy, DV protocols, immigration

Face Validity: The Pipeline Recovers Known Patterns

Gascón Transformation

LA County ideology score **tripled** under Gascón

Cohen's $d = 0.75$, $p < 0.001$

Geographic Clustering

- **Progressive:** Sacramento +78%, Yolo +56%, San Diego +50%
- **Traditional:** Stanislaus −34%, Placer −21%
- Bay Area variation: Santa Clara +0.84 vs Alameda −0.15

2020 Racial Justice Surge

Racial justice emphasis jumped **+30pp** in one year (12% → 42%), tracking the post-George Floyd moment precisely.

Documents with high racial justice emphasis are **4.6**× more likely to be progressive ($\chi^2 = 421$, $p < 0.001$).

Why This Matters

Pipeline confirms patterns any expert would expect—evidence of validity, not artifact.

Descriptive Analysis: Theory-Grounded Patterns

Extensive > Intensive Margin

Recent reforms disproportionately emphasize **who enters** the system (**33.9%** extensive lenient) over **how severely** people are treated (**22.6%** intensive lenient). Political economy logic: diversion/declination less visible to voters than sentencing leniency—safer reforms for DAs facing reelection.

Close Elections → Progressive Policy

Elections with margins $\leq$15pp produce **+31.2pp** more progressive policies ($p = 0.010$). Continuous: $r = -0.50$ between margin and ideology ($p = 0.009$). Provides credible first stage for instrumental variables designs.

Novel contributions: The extensive-over-intensive pattern has not been documented at this scale. The election-ideology link provides the first stage for causal designs.

The Variation Exists: Policy Shocks = Natural Experiments

What We Found

- **9** sharp policy disruptions (2020–23)
- **347** novel reform adoptions tracked
- Progressive docs: 18% → 56% in 4 years

Each Feature → a Causal Design

- 9 disruptions → **staggered DiD / event study**
- ∼5 close elections → **RDD first stage**
- DA transitions → **synthetic control**

The Key Insight

All four pillars of modern causal inference—DiD, RDD, synthetic control, and instrumental variables—are available in this data. No other prosecutorial dataset offers this.

What's Missing

Outcome linkage (Extension 2) and geographic scale (Extension 1) to power these designs. The policy variation is already in hand.

Pilot Results: The Signal Is Real

Finding	Estimate	Significance	Survives Controls?
Ideology ↔ Jail Pop Rate	$r = -0.222$	$p = 0.009$	**Yes** (year-demeaned)
Ideology ↔ Jail Admissions	$r = -0.221$	$p = 0.009$	**Yes** (year-demeaned)
Progressive vs Traditional	-68.5/100k	$d = -0.81$	**Yes**
LA Pretrial (Gascón DiD)	-32.1/100k	Pre-trend $p = 0.90$	**Cleanest**

What Works

All associations **survive year fixed effects**—not a COVID artifact. Pretrial detention shows the cleanest causal signal (pre-trend $p = 0.90$). These results motivate the full-scale designs on the next slides.

What Extensions Will Solve

Pilot N is small (34 CA counties, 9 years). National expansion + case-level outcome data will power the causal designs this variation makes possible.

The Payoff: 6+ Credible Causal Studies, Ready to Run

Policy Coded	Causal Question	Design	Outcome Data
Bail reform	Does reform reduce pretrial detention without increasing FTA?	Event study; DiD	Vera (done); CA DOJ
Diversion	Do diversion memos lower recidivism?	Staggered DiD	UniCourt; CJARS
Declination memos	Do "decline to prosecute" orders lower filings?	Synthetic control	CA DOJ (free)
Enhancement reform	Does curbing 3-strikes reduce sentence lengths?	Event study	CA Sentencing Comm.
Racial equity	Do equity directives reduce B/W disparities?	DiD; IV	CA DOJ by-race; CJARS
Close elections	Does a progressive DA *cause* lower incarceration?	RDD	Vera (done)

6 designs × multiple outcomes = a research agenda. The policy variation is in hand. Outcome linkage (Phase 1–2) unlocks all of them.

Deep Dive: Identifying the Causal Effect of Diversion Policies

The Question

Does adopting a diversion program **reduce recidivism** and **lower case volumes** without increasing public safety risk?

Identification Strategy

- **Treatment:** date a DA issues an internal diversion directive (coded by our pipeline)
- **Design:** staggered DiD across counties adopting at different times
- **Robustness:** synthetic control for Sacramento, Santa Clara

Outcome Data Needed

- **CJARS:** individual trajectories—does diversion reduce P(re-arrest)?
- **UniCourt:** case outcomes—do filing rates drop? Do plea rates change?
- **CA DOJ:** county-level prosecution counts (free, immediate)

Why Both

CJARS = **individual mechanism**. UniCourt = **case-level detail**. Together: "what happens to diverted people *and* to case flow."

Extension 1: Scale → Statistical Power for Causal Designs

Complete California

- Add **17 remaining counties** → full 58-county panel
- **Pre-2015 docs** → 10+ year pre-periods for event study
- **200-doc human validation** → publication-ready IRR

100 Largest US DA Offices

- **30–50 close elections** → statistically powered RDD
- Staggered treatment across states → robust DiD
- Cross-state variation → external validity

What Funding Unlocks

CA alone: ∼5 close elections (underpowered RDD). National expansion: **30–50 close elections**—enough for a credible regression discontinuity that can answer "does electing a progressive DA *cause* lower incarceration?"

Extension 2: Outcome Linkage → True Causal Estimates

Aggregate → County-Level Causal

Source	Causal Question It Answers
CA DOJ	Do filing rates drop after "decline to prosecute" memos?
Sentencing Comm.	Do enhancement filings fall after reform?
Vera (done)	DiD/event study on jail populations

Case-Level → Mechanism Evidence

Source	Causal Question It Answers
UniCourt	Do diversion memos change plea bargaining?
CJARS	Does a diversion memo reduce P(prosecution | arrest)?

What Funding Unlocks

Without outcome linkage, we have **policy variation without a dependent variable**. Each data source above converts one row of the "6+ causal studies" table from potential into a publishable result.

Value for Policymakers

- **DA accountability scorecards**—comparable measures of how each office's stated policies align with its goals
- **Rhetoric vs. practice**—does a "decline low-level drugs" memo actually reduce drug prosecutions?
- **Benchmarking tools**—a DA considering bail reform can see how similar policies performed elsewhere
- **Causal evidence on the progressive prosecutor model**—answering the central question: does it work?

Related Work

Felix Owusu's AV-funded project exploits two specific internal DA memos for a causal design. Our contribution scales this logic: **systematic measurement across 41+ offices and 1,865 documents**, creating variation needed for average causal effects and heterogeneity analysis.

Potential Funding Package

Phase	Components	Time	Budget	Causal Unlock	Deliverable
1	CA DOJ linkage; complete CA panel; human validation	6 mo	\$45–50k	DiD + synthetic control with validated measurement	First causal estimates of DA policy effects
2	100 US DA offices; national election DB; CJARS app	12–18 mo	\$100–150k	Powered RDD (30–50 elections); robust staggered DiD	National DA ideology database
3	Case-level data (UniCourt/CJARS)	6–12 mo	\$50–65k	IV + mechanism decomposition	Evidence of *how* policy changes courtrooms

Today: suggestive correlations (34-county pilot) → **Phase 1:** first credible causal estimates → **Phase 3:** full research program.

Why Fund This Now

Unique Advantages

- **Infrastructure is built**—pipeline, schema, analysis framework operational
- **Data window is closing**—ACLU PRA archive may not persist
- **No competitor**—no other systematic, AI-coded DA policy database exists
- **Pilot-tested**—demonstrated linkage to outcomes; identified the right methods

Policy Impact

- Answers the question **funders care about most**: do progressive reforms actually work?
- Enables **evidence-based evaluation** of DA accountability efforts
- Creates **benchmarking tools** for jurisdictions
- First **causal estimates** of prosecutorial policy effects

Let's Talk

41 counties · 1,865 documents · 6+ causal designs ready to run

The policy variation exists. The causal signal is real.
Fund the outcome linkage, and this becomes a research program.

Dvir Yogev
Post-Doctoral Researcher · UC Berkeley School of Law
dyo@berkeley.edu

github.com/dyo2112/prosecutor-policies-causal-inference

Full pilot report, figures, and replication code available in repository

Additional Detail

Overcoming Pilot Limitations

Limitation	Root Cause	Solution
COVID confounds DiD	All CA counties declined in 2020	Multi-state data; synthetic control
Parallel trends violated	LA converging pre-Gascón	Longer pre-period + case-level data
TWFE non-significant	34 counties, 9 years	National expansion increases N
Aggregate data too coarse	Cannot distinguish mechanisms	Case-level data from CJARS

Causal Identification Strategies

Design	Current Status	What's Needed	Extension
RDD	$\sim$5 close CA elections	30–50 close elections	Ext 1 (national)
Stacked DiD	9 disruptions; Vera only	More outcomes + transitions	Ext 1 + 2
Synthetic Control	Gascón pilot works	More outcome data	Ext 2
Event Study	6-year pre-period	Longer pre-period	Ext 1
IV	Margin $\rightarrow$ ideology ($r = -0.50$)	Outcome data for 2nd stage	Ext 2

Outcome Data Sources (Detail)

Source	Records	Timeline	Cost	Priority
CA DOJ OpenJustice	County-yr arrest, filing, incarceration	Immediate	Free	High
CA Sentencing Comm.	Enhancement filings, 3-strikes	1–2 mo	Free	High
Vera (done)	Quarterly jail pop and admissions	Done	Done	Done
UniCourt / PACER	Case dispositions, plea bargains	2–3 mo	\$30–50k	Very high
CJARS	Arrest $\rightarrow$ parole, 24 states	3–12 mo	\$5k	Highest

Case-level data is the gold standard: it links policy documents directly to line-prosecutor behavior.

Policy Disruption Detection: Methodology

Signal	Weight	Method
Ideology Velocity	30%	Rate of ideology change vs. prior 2-year baseline
Novelty Index	25%	Proportion of first-time policy types
Topic Shift	20%	Jensen-Shannon divergence of topic distributions
Margin Reversal	15%	Flips in extensive/intensive leniency direction
DA Transition	10%	New administration detection

Top disruptions: SF 2020 (Boudin, 0.572), LA 2021 (Gascón, 0.549), Sacramento 2022 (0.412)

Anticipated Objections (1/2)

"Policy changes aren't random. Where's the identification strategy?"

(1) **Close elections are as-good-as-random**—RDD at 50% cutoff; national expansion gives 30–50 races. (2) **Election margin is an IV**—$r = -0.50$ with ideology. (3) **Memo timing is plausibly exogenous**—the exact date a DA issues a specific memo has quasi-random administrative variation; staggered DiD exploits this. (4) **Synthetic control** needs parallel trends, not randomization—LA pilot: $p_{\text{pre-trend}} = 0.90$. (5) **Triangulation**: when all four designs agree, combined evidence is strong.

"Policy documents $\neq$ practice. How do you know memos change anything?"

That's why Extension 2 is Priority 1. We test whether "diversion-positive" memos predict more diversions. Even the stated-policy variation is *itself* a novel descriptive contribution.

Anticipated Objections (2/2)

"You only have California. Isn't this just a Blue State story?"

That's the point of Extension 1. CA has internal variation (Bay Area vs. Central Valley), but external validity requires the 100 largest DA offices nationally.

"AI coding isn't reliable enough for research."

Face validity is strong ($d = 0.75$ on Gascón, geographic clustering matches expectations). Phase 1 includes 200-doc human validation for gold-standard IRR. Pipeline cost: $80, runtime: 2 hours.