Essays on Local Governments: Finances and Crime

Tyler Ludwig
Great Falls, Montana

M.A. Economics, University of Virginia, 2018
B.A. Economics, University of Florida, 2017
B.S. Statistics, University of Florida, 2017

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Economics

University of Virginia
August, 2023

Leora Friedberg ______________________

Amalia Miller ______________________

Lee Lockwood ______________________

Rich Hynes ______________________

Abstract: Essays on Local Governments: Finances and Crime

I study local governments under fiscal stress and local elections of partisan politicians. In the first chapter, I explore municipal pensions. Most U.S. cities have defined-benefit pensions for their public workers, creating an obligation that exposes sponsoring cities to shortfall risk. Large funding gaps in recent years have required increased pension payments and generated fiscal stress for cities. To analyze the effect of "pension pressure", I assembled a novel dataset which captures the universe of cities and their pensions in California from 2003 to 2016. I focus on the changes in city unfunded liability contributions, which are plausibly exogenous to cities' year-to-year spending needs. Using a first-differenced empirical specification, I find that cities primarily target reductions in non-current expenses, specifically capital investment. I also show that cities cut payrolls and employment. Police employment declines have accompanying increases and decreases in crime and arrest rates, respectively. These estimates imply that pension pressure impairs local public service provision.

The second chapter, coauthored with Brett Fischer (California Policy Lab at UC Berkeley), considers elected district attorneys (DAs), who have wide discretion in criminal prosecution; yet, the extent to which DAs' politics shape their decision-making remains unclear. We evaluate the causal impact of DA partisan affiliation on prosecution rates, sentencing outcomes, and recidivism. Using quasi-random variation in DA partisanship stemming from close elections, we find that the marginal Democratic DA is 25 percent more likely to dismiss criminal cases than Republican counterparts, and 16 percent less likely to incarcerate defendants. Strikingly, defendants in jurisdictions with Democratic DAs are no more likely to reappear in a future criminal case, which is consistent with the notion that criminal prosecutions have limited deterrence or scarring effects. We find similar patterns using an alternative matching specification, suggesting our results capture the average effect of prosecutor partisanship. Our findings underscore the extent to which the punitiveness of the court system depends on the partisanship of local district attorneys.

In the third chapter, I examine the role of state oversight and interventions into local finances. I look at two cases, California and Ohio, which have different systems and contexts that allow for different empirical methodologies. Although some previous research suggests that there are economically large and statistically significant impacts on budgetary outcomes for local governments, my own analysis indicates that there is not enough evidence to reach definitive conclusions.

JEL Classifications: H7 K4 R5

Keywords: local government, city budget, pension, crime, district attorney, recidivism, fiscal monitoring, fiscal stress

Acknowledgements

I want to thank Leora Friedberg, who over the past many years has provided an abundance of advice and guidance. Even more importantly, Leora has keenly believed in my ability as an economist, which has heartened my pursuits as a researcher even when so many of them faltered. Further, I would also like to thank Amalia Miller, for her always useful perspective and economic insight; and Lee Lockwood for his rare talent to give such encouraging critiques of my work and presentations. More broadly, I have also greatly benefited from the support of the Department of Economics and its many members over the 6 years as I have progressed from preparing for core exams to writing research papers.

On a more personal note, I also want to show gratitude for Ishita Gambhir, for whom I do not have words to fully express how much her role in my life has meant to me. I have also been blessed to have Yooseon Hwang and others from within and without my cohort for their friendship and as sounding boards for ideas, economics-related and otherwise. Lastly, I would like to thank my family just for being there, even if my graduate studies have been somewhat shrouded from them.

Contents

Chapter 1

Pension Pressure: Impact of Public Pension Fund Liabilities on Cities

1.1 Introduction

In the United States, state and local public workers largely have their retirements secured by defined benefit (DB) pensions, guaranteeing them an annuity in retirement. Compared to the private sector where they are uncommon, 83% of full-time state and local workers participate in a DB pension plan.[1] The extent of these plans corresponds to a substantial obligation for the sponsoring government employers, and ultimately, their taxpayers: the overall present value of the benefits expected to be paid to these retired public workers is nearly $6 trillion.[2]

To meet this obligation, governments and their employees make yearly contributions to a retirement system, which then invests the contributions. Unlike with other types of

[1] "National Compensation Survey: Employee Benefits in the United States, March 2018," Employee Benefits Survey, U.S. Bureau of Labor Statistics.

[2] Author's calculations using the Public Plans Database developed by the Center for Retirement Research at Boston College. The figure reported is based on the plans' own actuarial assumptions, including long-term asset returns and retirement patterns. Other work, such as Novy-Marx and Rauh (2014), argue that the plans' should use a discount rate less than their assumed rates of return, which are generally around 7%, which would increase liabilities substantially.

plans, such as individual retirement accounts, governments are exposed both to investment volatility in their largely equity-based portfolios and to the shortfall risk of their invested contributions inadequately covering liabilities. Middling investment returns in the early 2000s were exacerbated by the Great Recession; in 2009, pensions lost nearly a quarter of the market value of their assets. This has created persistent asset-liability gaps measured in hundreds of billions dollars for local governments which are likely to deteriorate further in any near-term recession or low return investment environment. To shore up their funds, retirement systems have increased the contributions government employers must make. In total, state and local governments spent about 4.3% of their total $3.7 trillion expenditures in 2017 on pension contributions, up from an expenditure share less than half that size in 2002.[3] A few cities across the U.S. have even faltered under the resulting fiscal stress, with places like Stockton, California and Central Falls, Rhode Island filing for bankruptcy in the early 2010s. Altogether, the "pension pressure" placed on cities forms a large and lasting challenge to their fiscal health. I provide insight into the local impacts of short-term changes in pension pressure.

In this paper, I look at how cities accommodate changes in their required retirement contributions to DB pensions by altering their budgets and provision of local public services. The size of their obligations is driven by a stock of past decisions, rather than current choices; the onus comes not from new workers, for whom the present value of future pension disbursement is quite small, but from obligations to retired and near-retiring workers. Pension promises to workers are considered inalienable contracts, preventing cities from addressing pension pressure through a reduction of pension benefits for existing public employees.[4] For existing employees, this means that benefits can only become more generous, and reforms like higher retirement ages affect only new workers - reforms which are only long-term solutions rather than short-term relief.[5] Cities must resolve this pension pressure

[3]Author's calculations using Census of Governments and Annual Survey of Public Pensions.

[4]This "California Rule" interpretation of DB pension promises is followed by a number of states and has been established through over 70 years of court cases.

[5]CalPERS. "Vested Rights of CalPERS Members: Protecting the pension promises made to public em-

in their budgets now by raising taxes, cutting funding for services, or in rare cases, reneging on other debts. The choices cities make in response to pension pressure has implications for who bears the cost, as well as the short-run marginal value placed on different services and budget items.

I develop a novel data set covering California's cities from 2003 to 2016. The key components are city-level pension information derived from yearly actuarial valuation reports for over 1,000 pension plans and uniform financial reports on city budgets compiled by the state government. I focus on variation in cities' contributions to cover unfunded liabilities, which are mandatory payments resulting from funding gaps. I analyze how these contributions alter city budgets and outcomes using a first differences model capturing year-to-year changes, and provide estimates of the effects of pension pressure specifically and of mandatory expenditures more generally. I argue that these are plausibly causally identified given several key institutional details. First, compared to other portions of their required contributions, these payments are not based on the current size of the cities' payrolls. Second, the payments cannot be shirked in this sample, which includes only cities which have their assets managed by a state agency; even during its bankruptcy crisis when it reneged on other debts, Stockton made its contributions. Third, changes in contributions to unfunded liabilities stem primarily from pension asset shocks, which are further lagged two years when translated from actuarial valuation to the payments cities make; these contribution changes are unlikely to be associated with idiosyncratic changes in local economic conditions and tax bases. Further, cities do not appear to engage in much smoothing in anticipation of contribution changes.

In my preferred specification I find that, on average, when the city's unfunded liability cost – where the unfunded liability cost is the component of pension contributions that is required and arguably exogenous – increases by one dollar, current expenditures of California cities go up by $0.43. The cities' current expenditures consist of their employees'

ployees." July 2011. Web, accessed 31 Mar. 2019. This work can be found at: Link wages and benefits, private contracting, and other miscellaneous expenses. Spending on retirement reported by cities increases by \$0.81, which is partially offset a \$0.34 decrease in wages, suggesting that cities reduce employment, cut salaries, or both. Job cuts are one part of this response to pension pressure: police employment, which is on average about one quarter of total city employment, sees a loss of nearly 0.07 jobs per 100,000 city residents for every dollar increase in pension pressure. Combining this with the average 2005 to 2015 change in UAL cost of 34 and comparing with the average 2005 police employment of 211, this suggests an average long-term reduction of 2.3 (1.1%) paid police positions per 100,000 residents. If the marginal police officer provides social benefit through crime reduction, as in Mello (2019), the reduced public safety employment could lead to reductions in residents' welfare. I find evidence that crime rates rise, and the estimated direct costs of crime increase by around \$0.46 per capita for every dollar increase in the UAL cost.[6] Non-current spending, which includes investment in land, buildings, improvements, and equipment as well as debt payments, goes down by slightly more than a dollar. Non-current expenses are only partially funded from current revenues, with the rest funded through future revenues via debt. I find that cities reduce the level of debt that they hold, suggesting cities are compelled to service pension debts rather than debts for capital investment. These results are robust across a range of specifications, including those with untrended and trended city demographic controls, city-level trends, year fixed effects, and population weighting. The results imply that past public service provision in the form of city worker benefits are weighing on present, and through reduced capital investment, future public service provision.

The literature has said much on both the size of the pension problem and the political mechanisms that drive it. In 2009 near the financial crisis's height, the gap between assets and liabilities for state-level public pensions in the U.S. was calculated to be \$3.23 trillion when using market discount rates (Novy-Marx and Rauh, 2009). In both recent theory and empirical work, generous public pensions have been variously linked to public worker union

[6] To construct a cost-weighted measure of crime, I follow previous work and use \$67,794 and \$4,064 for the average weighted direct costs of violent and property crimes as estimated in Autor et al. (2017).

political clout (Kelley, 2014; Bouton et al., 2020), the ignorance of the median voter to complex pension issues (Glaeser and Ponzetto, 2014), poor local voter turnout (Trounstine, 2013), institutional constraints on debt and expenditure (Bouton et al., 2020; Glaeser, 2013), and political competition between parties for valuable union votes (Dippel, 2022; Bagchi, 2019).

However, less is known about the consequences of public worker pensions, especially their effects on cities.[7] At the state level, Shoag (2013) indicates that the investment returns of state-controlled pension plans significantly affects government spending, generating secondary effects on income and employment. Through a regression discontinuity design using the San Diego city boundary, MacKay (2014) finds that negative news about the city pension decreased housing prices. This paper contributes to the discussion by analyzing how cities alter spending and employment in response to pension debts.

I also contribute to a more general literature examining fiscal shocks and their effects on local government budgets. On the expenditure side – where my paper contributes – there are relatively few papers. Baicker (2004) finds that expensive capital crime trials lead to increased taxes and some decreases in police spending and investments. Similarly, I find that pension pressure sees cities reducing police employment and capital investment. On the revenue side, Shoag et al. (2019) shows that cities react to tax losses from large retail store closures by reducing police and administrative spending and raising revenue elsewhere. The cities in my sample also likely reduce their non-retirement public safety and administrative expenditures, but data limitations mean I cannot separate out the increase in spending in these categories from their contributions to their pensions. Similarly, reductions in local revenue induced by the Great Recession or by Chinese import exposure have been tied to reductions in spending on public goods like education, recreation, and waste management (Cromwell and Ihlanfeldt, 2015; Feler and Senses, 2017). Further, local governments facing

[7] A political science paper, Anzia (2022), surveys the issue using Comprehensive Annual Financial Reports (CAFR). Their findings complement mine by correlating total retirement spending with public-sector employment and collective bargaining.

reductions in military personnel presence after the Cold War reduced capital and increased debt (Komarek and Wagner, 2021). Natural disasters like California wildfires and Atlantic hurricanes also form fiscal shocks that have long-term consequences on municipal budgets through both the revenue and expenditure pathways (Liao and Kousky, 2022; Jerch et al., 2020). My results are overall similar but not identical to those in the literature: as in many papers on local fiscal stress, capital investments are the main target for cuts. I also find that wages decrease as benefit payments rise. Because I can go beyond data on spending patterns alone, I also show that cities cut workers from their payrolls – namely, police officers – with public safety consequences suggested by an increase in crime and its costs.

1.2 Background

1.2.1 Pension Accounting

In California, the vast majority of cities have DB pensions in contract with the California Public Employees' Retirement System (CalPERS). Most of the largest cities, like Los Angeles and San Diego, as well as some smaller cities like Alameda or Emeryville, maintain their own retirement systems in addition to or in place of contracting with CalPERS. CalPERS and the few independent retirement systems manage the cities' assets, meeting current and future obligations to members through a combination of investment returns and contributions. These contributions come from both working members and their employers, and cover two sources of cost; Figure 1.1 shows these costs. The first is the normal cost, which is intended to cover the present value of future benefits for each working member's additional year of service. Employee payments towards the normal cost are typically a set percentage of wages defined in the employees' contracts. Employers - that is, the cities - pay the rest of the normal cost not covered by employees. The total normal cost paid is relatively constant for cities. The second is the amortization of the unfunded actuarial liability (UAL), which is paid only by employers. My paper mainly focuses on city-year variation in contributions on the unfunded liability, as opposed to the contributions as a whole, in my empirical approach. Before describing it further, I formalize the contributions made by employers and employees in equations. Each year t, retirement systems assess each city i's pension plans, creating actuarial valuations that determine their required contributions, where the normal cost is split,

$$Contributions_{it}^{fixed} = NormalCost_{it} + UALPayment_{it}$$

$$NormalCost_{it} = NormalCost_{it}^{city} + NormalCost_{i}t^{employee}.$$

The UAL is a common summary measure of the funding gap formed by the difference between the value of assets and the actuarial accrued liabilities; that is, the UAL for city i can be expressed as $UAL_i = Liabilities_i - Assets_i$. Actuaries develop estimates for a pension plan's liabilities, which are the present value of expected future benefits from the prior service of retired and working plan members. The estimates are based on proprietary models, plan member demographics, and assumptions on mortality rates, service length, payroll growth, investment returns, among others. Reducing the complexities behind actuarial models for the purposes of providing background, liabilities can be written as the plan's estimated future benefits, discounted using the retirement system's assumed rate of return (ARR),

$$Liabilities_i = \sum_{t=1}^{\infty} \frac{Benefits_{it}}{(1 + ARR)^t}. \tag{1.1}$$

If there was limited investment return volatility and the actuarial assumptions made throughout the duration of the plan were correct, then normal cost contributions, which cover the liabilities accrued by workers within each year, would be enough for the invested contributions to exactly cover a plan's liabilities. But in reality assets fall below, and occasionally exceed, plan liabilities. Payments (credits in overfunded plans) towards the UAL are intended to close the gap between assets and liabilities, typically amortized over a 20- or 30-year period. Employers are solely responsible for this cost. Therefore, the presence of an unfunded liability adds to an employer's cost of maintaining a DB pension plan, and in many cases the UAL cost exceeds the entire normal cost. To give an example, Oakland's Miscellaneous Plan (which is for workers not in the emergency services) paid 22 million on the normal cost and 31 million on the unfunded liability in the fiscal year ending 2013.

Unlike the normal cost portion of the contributions, the UAL payment is prone to year-to-year fluctuations and changes over time. Primarily, changes in the UAL payment stem from asset markets. Volatility in the performance of the pension plan's investments means volatility in the size of the funding gap, changes which must be amortized.[8] Other factors, like changes in the assumed rate of return (ARR), which is the discount rate for future benefit payments, play a lesser role. Since they are responsible for its payment, employers bear the entirety of the UAL payment volatility. Moreover, employer contributions typically exceed employee contributions.[9] It is not possible for employers to manage the stresses placed on their budgets by their changing contributions through nonpayment or underpayment: CalPERS imposes large fines on contracting cities which fail to pay, essentially precluding this behavior from cities.

A few additional features of pensions are useful in analysis. First, shocks to pension assets are unlikely to be associated with local economic conditions, since pensions hold diversified portfolios heavily favoring global funds. As of June 30, 2016, CalPERS' Public Employees' Retirement Fund held 51.9% and 20.3% of its total investments in global equity and global fixed income, respectively (CalPERS, 2016) (CalPERS, 2016).[10] Second, there is a lag between actuarial valuations and the contributions determined from them. The CalPERS system has a two-year lag: for example, a given plan's actuarial valuation at the end of fiscal year 2012 determines the contributions for the fiscal year ending 2015. The other systems in California have one-year lags. This lag is useful empirically, since it reduces the

[8]Many pension plans dampen volatility through a variety of mechanisms. One method is smoothing market gains or losses across a number of years, called a smoothing period, to generate the "actuarial value of assets". For instance, a 10% decrease in market value of assets for a system with a four year smoothing period would only register as a 2.5% decrease in actuarial value in that year.

[9]E.g. for Anaheim's Miscellaneous plan FYE 2008, the city paid 14.953% and employees paid 8% of payroll.

[10]The next three largest asset classes are real assets (10.8%), private equity (9.0%), and inflation assets (6.0%).

simultaneity of the effects of macroeconomic shocks on pension contributions and on local tax revenue.[11] In my empirical approach, I use city-year variation in contributions on the unfunded liability, rather than variation in contributions from normal costs, which depend on the current size and composition of a city's workforce.

1.2.2 City Budgets

Compared to other layers of government, urban public finance has some unique characteristics. In the United States in general, and California specifically, cities have limited independence; their budgetary activities are ultimately constrained by their respective state governments (Glaeser, 2013). The California Constitution proscribes local governments from incurring any debts greater than their revenues, and thus imposes balanced budgets on cities. There are a few exceptions to this rule. Mainly, cities can issue municipal bonds to finance capital projects, but only with two-thirds voter approval. Changes in tax rates similarly require a super-majority vote. Cities also face caps in the form of the "Gann Limit", which like Proposition 13 is another legacy of the 1970s tax revolt. It limits the growth in taxes and expenditures based on population and income growth, though these do not seem be binding constraints (Kousser et al., 2008). Pension pressure, in the form of rising retirement expenses, comes into this setting and forces budget reallocations.

For exposition purposes, I present a model of a simple budget with these institutional details in mind. Each year a municipal government must create a balanced budget such that,

$$Expenses_t = Taxes_t + OtherRevenue_t + CapitalDebt_t - FundChange_t$$

with expenses further dividing into non-retirement current expenses, non-current expenses (e.g. investments into infrastructure), and retirement contributions,

$$Expenses_t = CurrentExpenses_t + NonCurrentExpenses_t + Contributions_t^{fixed}$$

where retirement expenses are externally determined (although changing over time) due to legal restrictions obligating the city to pay. Cities can meet their rising pension costs by either

[11] Although not shown in this paper, controlling for changes in tax revenues does not qualitatively alter the estimates to be reported in Section 5.

raising more money or dropping expenditures elsewhere; they can choose to cut services, raise taxes, take on debt, use up government funds, or some combination of these. From a simple economic model incorporating the incentives of cities and their politicians and using the above budget constraint, one can hypothesize that their specific choice will be guided by a combination of desire for reelection and maximizing resident happiness. I provide insight into these choices and as a result of pension pressure.

1.3 Data

1.3.1 Overview

I focus on California and its cities in my analysis. Cities are important entities in California, and on average spend over double per person than the state government spends.[12] Incorporated municipalities also contain nearly 80% of the state's residents. Figure 1.2 shows the market value of assets and accrued liability for the state's municipal pensions.[13] Statewide, city pensions were nearly fully funded in the years leading up to 2009, when pensions lost around a quarter of their investments. Despite some investment returns in the ensuing years, pension assets have failed to again meet the liabilities; in 2013's valuations, city pensions were unfunded by around $35 billion dollars, or about $1,100 per Californian living in cities. As they are based largely on the size of the gap, contributions have correspondingly changed, and range between around 0% to 10% of total city expenditure.

1.3.2 City Finance Data

To explore this topic, I developed a rich, novel panel of financial reports and pension funds covering all of the cities in California from 2003 to 2016. Data from the California State Controller provides detailed fiscal information on the nearly 482 cities in California. Each

[12] See: https://www.ppic.org/publication/the-state-local-fiscal-relationship/

[13] Here the liability is the Entry Age Normal Accrued Liability. As mentioned previously, the discount rate for calculating liabilities is higher than the market discount rate more typically used for present value.

fiscal year, all California cities are required by law to submit a Financial Transactions Report (FTR) according to a uniform classification system.[14] Among the information collected are (1) expenses for total wages, retirement, private contracting, and other total costs, (2) expenses for specific services such as general government, public safety, culture and leisure, and health, and (3) fund management in the form of debt service and issuance of long term debt, and (4) emergency services employment. Table 1 provides definitions for categorical expenditures and examples of what each includes. Uniformity and completeness allows a rare view into U.S. cities not seen in other data sets, and allows comparisons across cities not possible otherwise. For reference, one frequently employed dataset is the U.S. Census of Governments, which has gathered financial information from all levels of American government since 1957. However, the Census of Governments is limited by its relatively infrequent five-year recurrence, which cannot be remedied by the small and changing sample of the intercensal Annual Survey of State and Local Government Finances. I discuss comparisons between the FTRs and these other city finance data sources in 1.10

1.3.3 Pension Data

I pair this with city-level information on pension funds. For the cities I gathered data on plan standing, payroll, contributions, numbers and types of members, and other details from a number of sources. My primary source on pensions were actuarial valuation reports from CalPERS, which is by far the largest provider of public worker plans in California. Every fiscal year, CalPERS sends out actuarial valuations to each contracting public agency, which inform them of their plan's standing; I attained these via a public records request. All in all, there were a little over 17,000 actuarial valuations, each representing a different plan-year observation from the 15 years of interest. These valuations range from 10 to 80 pages and consequently differ considerably in their contents, both across time and between

[14] California State Controller's Office, Nov. 2018. "Cities Financial Transactions Report Instructions." Web, accessed 6 Apr. 2019. These instructions are available at: https://www.sco.ca.gov/Files-ARD-Local/LocRep/Cities%20FTR%20Instructions.pdf

plans in the same year.[15] To help give a sense of what these reports look like, I present two relevant pages from the Annual Valuation Report for Riverside, CA for 2005 at the end of 1.11 I scraped these valuations for pertinent information on the standing of municipal plans. A few cities have some or all of their pensions serviced by a non-CalPERS system, such as San Francisco by the San Francisco Employees' Retirement System. Although I don't use non-CalPERS cities in most of my empirical analysis, I am also able to scrape the plan information for many of these using actuarial valuations sourced from their website or by request. For the few pension-years which escaped this effort, I supplemented with the California State Controller's Public Retirement Systems Financial Data, which stem from yearly obligatory standardized reports like the city FTRs. From these sources, I aggregated all pension plans to the city level.[16]

I avoid using cities that have non-CalPERS plans because instead of state management they are controlled locally. The non-CalPERS retirement systems are generally city departments. For example, the Los Angeles City Employees' Retirement System is part of the city itself. It is plausible that cities with their own retirement system would be able to influence the pension's choice of investments, actuarial assumptions, or most importantly, the amount of contributions requested by the valuation or the amount paid by the city (if they choose to shirk). The consequent concern for possible endogeneity is not without precedent. For instance, in the early 2000s San Diego altered pension policy in order to avoid increased contribution payments resulting from investment losses (MacKay, 2014). Collectively, the cities that do remain after sample restrictions had 22.5 million residents in 2016.

Table 2 presents summary statistics on city expenditures for the CalPERS sample, including the normal and unfunded liability costs sourced from the annual valuations.

[15] These differences in contents stem from whether the plan was in a risk pool, which I describe further in 1.11

[16] Cities tend to have multiple plans. Within a city the different plans are divided into the categories of safety (e.g. firefighters and police) and miscellaneous, and then further by generosity (for instance, 2% at 60 versus 3% at 50). In 2015, the City of Oakland had two independently administered plans, and two CalPERS administered plans. Other cities have even more plans: Laguna Beach had nine plans through CalPERS in 2015.

Expenditures are shown in both inflation-adjusted per-capita dollars and shares of total expenditures. The table is broken into two cross-sections from the fiscal years ending 2005 and 2015, allowing some comparisons over time. One takeaway is that pension expenditure, on average, rises from about 11.8% to about 22.5% of city wages between the two years; UAL costs alone rise from about 2.2% to 10.8% of city wages. Another takeaway is that total expenditures are split about 80-20 between current and non-current expenditures, where non-current spending mostly consists of capital investments.

1.3.4 Additional Data

A handful of other sources supplement the pension and city financial data in order to provide a more complete picture of the region's cities. To transform variables to their real per capita equivalents, I use city population estimates and the fiscal year average California CPI from the California Department of Finance. The latter of these is constructed using a population-weighted average of the US Bureau of Labor Statistics CPIs for California locations.

To adjust for differences in municipal demographic characteristics, I also collect data on cities from the National Historical Geographic Information System (NHGIS).[17] The NHGIS ties Decennial Census information to cities using historical city shapefiles. This provides city-level data which is otherwise sparse. Unfortunately, intercensal demographic data is either limited, like to population (as used in this paper), or at county-level and above. The demographic and economic controls from the year 1990 are used to form a baseline prior to the start of the data series in 2003. Specifically, I use as controls the proportions of city residents who are white, black, Asian, Hispanic, under 25 years of age, and over 65 years of age; the proportion of households which are home-owners; and the median home price and median rent. In some empirical specifications I additionally allow the baseline controls to have changing effects over time by interacting each with a linear time trend.

[17] Source: Manson and Ruggles (2022)

Police employment data are from the FBI's Uniform Crime Reporting Data System (UCR). I gathered the agency-level Law Enforcement Officers Killed in Action (LEOKA) files for 2002–2015.[18] The LEOKA files report counts of police officers and civilian employees at each agency. I merged agencies to cities based on name. Not all cities have their own police agency: only about three-quarters of the cities in my sample do. My analysis of public safety employment outcomes is correspondingly restricted to the cities with independent police services. The summary statistics for the LEOKA-sourced police employment counts, along with the FTR-sourced fire and EMS employment counts, are shown in Table 3. I also acquired crime and arrest counts for index offenses from the UCR Return A files. I aggregate the crime counts to the agency-fiscal year level using the monthly data from 2002-2016. To account for record errors and outliers known to be issues in the UCR data, I implement a cleaning procedure similar to that documented in Mello (2019).[19] I transform these into crime and arrest rates per 100,000 residents by scaling them using city population. To construct a cost-weighted measure of crime, I follow previous work and use $67,794 and $4,064 for the average weighted direct costs of violent and property crimes as estimated in Autor et al. (2017).

1.4 Strategy

In my study, I aim to recover the effects of spending towards public retirement funds on budgets and public service provision for sponsoring cities in California. I discuss details that help inform my choice of model and present my empirical specification in this section.

[18]LEOKA data sourced from Kaplan (2021). Since LEOKA employment is reported in October in each calendar year, I assign them to the fiscal year ending in June the following calendar year.

[19]One key aspect of cleaning the data involves using local linear regressions to identify extreme observations relative to their nearest neighbors. If the changes in a count exceeds a threshold, defined by city population groups, then that count is set to missing and, if possible, overwritten by backwards filling, forwards filling, or interpolation to maintain panel integrity.

1.4.1 Setup

To expand on information provided so far, a few facts are useful in my analysis. First, shocks to pension assets are unlikely to be associated with local economic conditions, which would bias my estimates. Based on twenty states, Shoag (2011) finds that pension plans on average only over-allocate within their state with 0.31% of their portfolio relative to the share that would be allocated if the plan invested only in the Standard and Poor's 500 index; CalPERS specifically, the primary pension plan in this paper, is shown to have an in-state bias of 0.38%. Brown et al. (2015) finds a higher average bias at 4.1% of portfolios, based on 27 state plans for quarters in which they self-managed asset allocation, though this too does not present much threat to analysis. Further, my analysis concerns local governments whose pension assets are held outside of their control at the state-level. This advantage is provided by my novel data on municipal DB pensions which are contracted with a state retirement system; here, CalPERS. Second, there is a lag between actuarial valuations and the contributions determined from them, which reduces the simultaneity of the effects of macroeconomic shocks on both pension contributions and local tax revenue. In my regressions, I assume that there is no correlation in each differenced error with its differenced UAL payment, an assumption which seems more plausible given the institutional lag separating outcome and the regressor of interest temporally. Third, California municipal budgets have several constraints; most importantly, the balanced budget requirement means yearly changes in pressure affect the city within the year, rather than having a delayed effect in future years.

Fourth, retirement expenses vary drastically across both time and cities, allowing sufficient variation for identification. Over time, retirement expenses can change fairly drastically: city contributions to Los Angeles' three retirement plans increased by nearly a factor of 6 from 2003 to 2016. They also differ between cities based on generosity and funding history: different cities hold different stocks of pension assets due to historical choices. Therefore even though all cities experience similar percentage returns on investments, the associated gains and losses are based on an interaction with the city-specific asset stock. My analysis focuses on using these plausibly exogenous investment shocks interacting with city pension assets and their effects on city retirement expenses. In Figures 1.3 and 1.4 I show the between-city variation in pension pressure from 2005 to 2015. The former map focuses on the cities in the Bay Area and Central Valley, while that latter focuses on those of Southern California. These demonstrate that some cities have had little change in their real per-capita expenditure on pension expenditures through the decade including the Recession, while others have come under budgetary pressure. I also show that year-to-year changes in the UAL cost are different between cities across time in Figure 1.5. The plot further shows that changes in the UAL cost did not exclusively occur at the onset of the Great Recession. Instead, they occurred throughout the sample period, and the largest average increase preceded the Recession. Altogether, the variation thus demonstrated provides the basis for an empirical strategy using first-differences.

1.4.2 Specification

In order to uncover the effect of pension pressure on city behavior, I use the city's contribution based on only the unfunded liability. Within my context, cities must make mandatory unfunded liability contributions as determined by the state agency CalPERS. Cities may only influence the size of the contributions over the long term: for instance, by making enduring reductions in their covered payrolls. In the short term, cities cannot meaningfully influence these payments, and must instead adjust in other ways.

I take advantage of the short term inability to control using first-differenced panels of California's cities and pensions. For each city c at time t, I define

$$\Delta UALPayment_{c,t} = UALPayment_{c,t} - UALPayment_{c,t-1};$$

that is, as the change in the payment on the unfunded accrued liability (UAL) of the city's pensions relative to the level in the previous year. The change in the UAL contribution is measured in real per-capita 2016 dollars using yearly population estimates, as are other variables where appropriate. Similarly, $\Delta Y_{c,t}$ is the year-on-year change in an outcome of interest, like safety expenditure or employment. An empirical model of their relationship is

$$\Delta Y_{c,t} = \beta \Delta UALPayment_{c,t} + \gamma X_{c,t} + \epsilon_{c,t}, \tag{1.2}$$

where the coefficient β measures the contemporaneous one-year change in the outcome of interest from changes in the UAL cost. Further, $X_{c,t}$ is a vector of covariates, discussed further below; $\epsilon_{c,t}$ is the error term, clustered by city; and regressions are weighted by the city's average population during the sample period, fiscal year end 2003 through 2016, to enable β to better capture the average experience of cities' residents.

As with unit fixed-effects, first-differences removes any bias stemming from time-invariant differences between cities. The small affluent and urban city of Beverley Hills contrasts the sprawling cheap and agricultural city of Modesto in ways that likely influence both the levels in the outcome, like culture and recreation expenditure, and the size of the unfunded liability, like the extent of their civil service. This allows better cross-city comparisons, and for β to represent the effect of per-capita intensity of pension pressure on per-capita outcomes. Unlike the fixed-effect model, though, first-differences focuses on short term, year-on-year changes rather than deviations from the long term mean in each panel. A first-differences model reduces concern that estimates are distorted by cities making long-term choices that change both their unfunded liability payments and outcomes of interest. First-differences does reduce the effective number of observations, though, since any stand-alone missing values in the mildly unbalanced city panel lead to two missing differences. For this reason, I provide the number of observations in brackets for each estimate in the tables described in the results section.

One challenge to β measuring a causal effect arises if cities' differential "treatment" – that is, $\Delta UALPayment_{c,t}$ – is instead correlated with differential trends across cities, leading to spurious estimates. I seek to address this concern in several ways. In addition to including the year-on-year change in the log of population, $X_{c,t}$ contains a vector of baseline city demographic variables from the year 1990, over a decade from the beginning of the sample. The city demographic variables included are: the proportions of city residents who are white, black, Asian, and Hispanic; the proportions of residents who are below the age of 25 and over the age of 64; the proportion of residents who are homeowners; and the city's median home price and rent. Inclusion of these covariates helps account for any potential long term secular trends associated with city characteristics; for instance, if cities with a higher proportion of homeowners have had smaller expansions in city expenses.

To address possible empirical concerns, check for robustness, and examine heterogeneity, I modify the baseline first-differenced model in a few ways. For brevity I include all of these modifications in the following equation, though each are added in to the model iteratively in the results. The model inclusive of all modifications is

$$\Delta Y_{c,t} = \beta \Delta UALPayment_{c,t} + \gamma X_{c,t} + \eta X_{c,t} t + \phi_c + \rho_t + \epsilon_{c,t}. \tag{1.3}$$

First, I allow baseline city characteristics to have linearly time-varying impacts on year-on-year changes in outcomes by adding $X_{c,t}t$. Second, I include a city fixed-effect ϕ_c. In the context of a first-differenced model, the inclusion of a unit fixed-effect controls for unit-level linear time trends. Thus, the city fixed-effect adds a more stringent control to abate any contamination of estimates by secular trends. For instance, El Segundo, a small city in Los Angeles county with both an agglomeration of aerospace industries and large per capita pension debts may have had unique budget impacts from changes in national aerospace spending. Third, I add a year fixed-effect ρ_t to capture state-level shocks. While the institutional lag between actuarial valuations and the contributions determined from them removes the simultaneity of effects of macroeconomic shocks on pension contributions and city finances, national and state level policy, e.g. an expansion of the California government's grants to cities. Fourth, I present unweighted regressions, allowing each city to enter equally in the regressions. Therefore – after restricting my sample to cities with CalPERS plans only – the largest city, Long Beach, with a population around 450,000, enters the regression with the same weight as the smallest, Etna, with around 750 residents. The resulting unweighted estimates provide the average experience of cities, rather than the average experience of city residents.

1.4.3 Retirement Spending and Persistence

Before proceeding to a discussion of results, I first validate the connection between the UAL payment recorded in the actuarial valuations and the retirement spending recorded in the Financial Transaction Reports (FTRs). In Table 4, and in the tables on city spending and other outcomes to follow, I show estimates from five empirical specifications. Each column contains estimates produced from a different model. The first corresponds to Equation 1.2. The second, third, and fourth columns' estimates differ from column 1 through the iterative addition of trended baseline covariates, city fixed effects, and year fixed effects. The fifth column repeats column 1 with an unweighted specification to explore heterogeneity by city size through comparison to the other models' estimates. In Panel A, I perform a "sanity check" on the connection between my two primary data sources, CalPERS actuarial valuations and city Financial Transaction Reports (FTRs), each of which each have information on city retirement spending. I show that a one dollar change in the valuation-derived UAL payment is strongly linked with an increase in the retirement spending reported in the FTRs. The estimate in column 3, which includes city fixed effects, is an highly significant increase of \$0.81, with similar estimates in the other columns. I provide two comments on these point estimates. First, a 1:1 change is within the 95% confidence interval for the majority of models. Second, cities report in their FTRs any contribution to retirement funds, including both their contributions towards DB pension liabilities accrued in the current year (normal cost) and contributions towards the funding gap (UAL cost). The reduction in city payrolls that I find would simultaneously reduce the normal cost portion of the city's total retirement costs and thus pull the estimated relationship away from 1. In Panel B, I assess the significance of once, twice, and thrice lagged differences in the UAL cost on city retirement costs. Unlike the unlagged first-difference, these do not have a consistently significant association with retirement spending across the five models. This suggests that my sample's cities are not simply experiencing persistent changes in their UAL cost, and that first-differences is a reasonable model for this dynamic panel.

1.5 Results

In this section, I study contributions towards unfunded liabilities in municipal DB pensions, concentrating on their impacts on city spending, public safety employment, and crime. Throughout, I present estimates of β in Equation 1.2, which captures the plausibly causal short term responses cities have to changes in mandatory, externally determined expenditures. Baseline estimates of the effects on each outcome of interest are accompanied by those from models incorporating additional covariates, fixed effects, and weighting differences to demonstrate robustness to increasingly stringent controls and to examine heterogeneity.

1.5.1 Budget

I first estimate the effects on non-categorical city spending outcomes – that is, those that are not for the specific uses described in Table 1 like community development. These are reported in Table 5. Both the outcome variables and the cost associated with the unfunded actuarial liability (UAL) are presented in terms of the change in per capita, 2016 dollars. Outcomes in this table and the following tables are shown as rows, where the estimates displayed are the coefficient β on the term $\Delta UALPayment_{c,t}$.

A one dollar change in pension pressure does not alter total expenditure. Instead, cities increase current spending - which includes the UAL cost - and decrease non-current spending. Breaking down current spending, the total retirement spending reported by cities goes up by around \$0.80 across all specifications, as described previously. Cities' wage expenditure decreases around 33 cents for every dollar increase in the UAL cost. The effect on wages is highly statistically significant. This is not effectively a change in the price of city workers, as would be the case with a change in the normal cost, which covers current workers' pension liabilities accrued that year and increases with additional workers or higher wages. The amount of the unfunded liability is not dependent on current worker pay, so cannot explain the change in wages. Instead, one possible explanation is that if cities combine wages and retirement costs in their budgets as worker compensation, and there are some frictions in adjusting the budget allocated to compensation, a reduction in wages is a credible response to higher retirement costs. Breaking down non-current spending into capital and debt service, I find that cities reduce their spending on capital investments like buildings and equipment. The reductions are large - in column 4's model with city and year fixed effects, for each additional real per capita dollar spent on servicing the UAL cost, cities significantly spend \$1.14 less on non-current expenses and \$0.84 less on capital investment specifically. Note that investment is only partially funded from current revenues, with the rest funded through future revenues via debt. These are not unlike the findings in other research on local fiscal shocks; for example, Cromwell and Ihlanfeldt (2015), where capital investment receives the majority of the budget cut.

The Financial Transaction Reports also categorize expenses according to their use in providing city services. In Table 6, I examine the effect of the unfunded liability on current categorical expenditures; namely, (1) health, (2) safety, (3) general government, (4) transportation, (5) community development, and (6) culture/recreation. City governments vary in terms of which services and how they provide them, but these categories summarize the diverse set of services cities provide. Also, by aggregating over the many variables the financial transaction reports provide, I minimize the issue of multiple comparisons (e.g., looking at individual items like sewer capital expenses). Current spending on safety and general government, two categories with the most pensioned workers, rise with changes in the pension costs consistently in each specification. In the model with city fixed effects in Column 3, general government and safety spending increase by $0.38 and $0.32, respectively, with each dollar increase in the UAL cost. I assume the rises in spending in these categories are due to their pensions, which are typically the largest and most generous of all public workers in cities; that is, retirement expenditures inflate the cost of providing these services, as cities spend money to make up for past generosity and under-performing pension investments. Unfortunately, the Financial Transaction Reports do not separate categorical expenditures into wages, retirement, and other expenses. I instead look at employment below to translate these spending pattern changes into employment changes. Culture expense also rises significantly in some specifications, though the negative point estimate in the unweighted regression in column 5 indicates heterogeneity by city population; that is, that only large cities see increases in culture and recreation. One possible insight into the difference is that the median city with an average population under 10,000, a group which forms about one-fifth of the sample, spend about one-third less in terms of shares of total expenditure than the overall median city (5.0% versus 7.7% of total expenditure). Thus, these estimates suggest that pension pressure inflates the cost of city service provision, but data limitations mean that we cannot conclude whether service quality is affected or not from these estimates alone.

1.5.2 Safety Employment and Crime

To gain further insight into service provision, I look at public safety employment for cities which directly employ paid police, firefighters, and emergency medical staff.[20] In Table 7 I show that any rise in spending on safety is not driven by a concurrent rise in employment. In fact, the number of paid police personnel per 100,000 residents decreases

[20]Not all cities directly have paid position. For instance, police services may be contracted out to the county sheriff's department, firefighting departments may be volunteer, and emergency medical technicians may be employed privately by ambulance services and hospitals.

by around 0.07 per 100,000 residents as pension pressure increases, a reduction which is split between officers and civilian employees. There are negligible changes in firefighters and emergency medical technicians. The average 2005 to 2015 change in UAL cost is 34 and the average police employment was around 211 in 2005. Combined with the estimate, back-of-the-envelope calculations indicate an average long-term reduction of 2.3 (1.1%) paid police positions per 100,000 residents. In magnitude, this is around 5% the magnitude of the estimated per capita effect of the police employment grant studied in Mello (2019). So even as public safety takes up a larger portion of the budget (in an accounting sense) at the detriment of other categories, taxpayer money is going to paying down previously generous service rather than maintain or improve current service.

The documented police employment decreases encourage exploring their public safety effects. I use the crime and arrest rates as reported in the city-level FBI Universal Crime Reporting Return A data from 2003-2016. More specifically, Table 8 includes as outcomes the year-to-year changes in both total property and violent crime per 100,000 residents. I consider the model with city fixed effects in column 3, which accounts for city-level trends in crime and arrest rates. For each one dollar increase in the city's unfunded liability cost, the property crime rate significantly increases by 3.3 (0.10% of the mean) and the violent crime rate increases by 0.56 (0.05% of the mean). Furthermore, assuming the costs of the average violent and average property crimes are $67,794 and $4,064, the estimated direct costs of crime increase by about $0.52 per capita, and are significant in both statistical and economic senses. I do not find any statistically significant changes in the arrest rates per capita across my models. However, the change in the property crime clearance rate - defined as the ratio of arrests to crime within a fiscal year - is significantly related to the change in the unfunded liability contribution in all but the model with city and year fixed effects. From column 3, I estimate that a one dollar increase corresponds to a 0.024 percentage point (0.16% of the mean) reduction in the clearance rate.

The results should not necessarily be interpreted as stemming only from policing changes, however. For instance, while the reduced police employment documented previously is possibly an important channel for changes in the crime rates, an idea which is bolstered by the simultaneous reduction in arrests for property crimes, the estimates are ultimately reduced form. Other changes, like those in non-public safety services, could also help explain the results. Still, I further compute IV estimates for police employment on the costs of crime victimization, using a two-stage specification of the form,

$$\Delta Police_{c,t-1} = \theta \Delta UALPayment_{c,t-1} + \gamma X_{c,t-1} + \eta X_{c,t-1}(t-1) + \phi_c + \epsilon_{c,t-1} \tag{1.4}$$

$$\Delta CostCrime_{c,t} = \beta \Delta \widehat{Police_{c,t-1}} + \nu X_{c,t} + \rho X_{c,t} t + \tau_c + \sigma_{c,t} \tag{1.5}$$

where police employment and UAL payments are once-lagged in reference to the costs of crime victimization. I estimate a significant reduction in \$56 per capita for each additional police employee per 10,000 residents. This can most directly be compared to Mello's (2018) estimate of a reduction of \$35 per capita, generated by using the federal police employment grants as an instrument. Though my estimate is sensitive to model choices in both magnitude and significance, it helps support the hypothesis that the change in police officers is the main channel for the crime increase. Together, my results provide suggestive evidence that the strain of pension pressure on city fiscal resources leads to an increase in crime.

1.5.3 Other Outcomes

I also look at how municipal debt is influenced by pension liabilities. I present four different outcomes in Table 9 - (1) the number of outstanding debts, (2) whether new debt is issued, (3) the amount of debt issued, and (4) the outstanding principal. These outcomes are worth exploring since investment, which is largely funded through debt, is the main lever cities use in response to pension pressure as reported in Appendix Table 5. Cities may seek debt for other reasons as well. The model in Bouton et al. (2016) supposes that local governments simultaneously choose the levels of debt and retirement entitlements for the benefit of interest groups, and shows that they may act as substitutes or complements depending on the institutional setting. Anecdotally, a few cities in California, such as Healdsburg, Monrovia, and Pleasant Hill, explicitly took loans or issued bonds to pay down their unfunded liabilities, and therefore reduce their yearly pension cost.[21][22] The number of outstanding debts does not change. Looking at the other three outcomes in the Table 9, I find that the cities issue new debt but the amount of outstanding principal still declines. Thus, the reduction of capital investment has an associated reduction in principal, outweighing any other debt-seeking.

1.5.4 Additional Robustness and Heterogeneity of Key Results

Lastly, I examine the robustness and heterogeneity of the results. In Appendix Table 11, I provide estimates of the key elements from the finance, police employment, and crime tables. Column 1 reproduces earlier estimates from a model in first-differences with a city fixed-effect. Column 2 shows that all of the key results are robust to trimming the top and bottom 5% of observed changes in the UAL cost. In columns 3 and 4, I restrict the sample to cities which held above and below median assets per capita, respectively, in 2003. I find that the above median assets cities have results that are again comparable to column 1 across the results. For below median assets cities I find that the estimated effect on total city payrolls is comparable but several other estimates lose their significance and the estimated effect on non-current spending is near-zero. Note that the connection between changes in employer UAL cost, as reported in the cities' actuarial valuations of their pensions, and changes in the retirement spending, as reported in the cities' financial transaction reports, is smaller with a retirement spending increase of $0.61 for every dollar in UAL costs. Finding blunted effects across the results in the below median asset cities is not surprising since the variation in UAL costs is correspondingly reduced.

[21] See the following links: Healdsburg, Monrovia, and Pleasant Hill

[22] These cities had high implied rates of interest on the unfunded liability (the actuarially assumed rate of return is around 7.5% at this time in CalPERS), so they refinanced using cheaper debt.

In columns 5 and 6 I show the results of non-differenced fixed-effect specifications. That is, specifications of the form

$$Y_{c,t} = \beta UALPayment_{c,t} + \gamma X_{c,t} + \eta X_{c,t} t + \phi_c + \rho_t + \epsilon_{c,t}. \tag{1.6}$$

As mentioned in the research design section, these no longer capture the within-year responses of cities to changes in pension pressure, and instead show longer term outcomes which likely reflect endogenous choices by the cities. Still, with the inclusion of a linear city time trend to help capture patterns in city spending and outcomes, the estimates in column 6 are qualitatively similar to those in column 1. I find comparable reductions in wages, non-current spending, and police employment, and an increase in the estimated costs of crime victimization, though estimates are generally noisier.

1.6 Conclusion

The growth of costs for public worker pensions represents a serious challenge for local governments and taxpayers, and has been an especially keen issue in the aftermath of the Great Recession. In 2017, over 99% of cities with pensions in California had funding gaps in their retirement plans, and half of these had unfunded liabilities over $1,000 per resident. The literature has concentrated on documenting the scope of the problem, as well as theorizing about its political and economic origins. Yet it has been relatively silent about how cities are affected, and how they choose to confront this pension pressure. One reason is that researchers have lacked quality data linking municipal pensions to their cities. I bridge this gap by developing a novel data set. I link rich city financial data to researcher-collected information from the state retirement system to provide the universe of municipal pensions and cities in California. Using the resulting panel and first-differenced models, as well as institutional information, I estimate plausibly causal effects. I find that in response to growing unfunded pension liabilities cities reduce spending on non-current expenses, specifically capital investment, and wages as benefit payments rise. I also show that beyond just changes in spending patterns, cities cut workers from their payrolls. These cuts include police officers, which could have public safety consequences including changes in crime rates. This implies that taxpayers are getting worse public services for their dollar, with contributions on pension debts displacing spending in cities.

Future research should continue to examine the public pensions of local governments, as much of the research has focused on the state level. The lowest layers of governments – cities, counties, school districts, water districts, fire districts, and more – provide public goods and services which directly impact the everyday lives of taxpayers. Most of these have defined pensions for their public workers. Collectively, the debts hidden in their public worker pensions will present an important and growing threat to the fiscal health of local governments, impacting the amenities they provide and the revenue they demand.

1.7 Figures

Figure 1.1: Diagram of Pension Payments

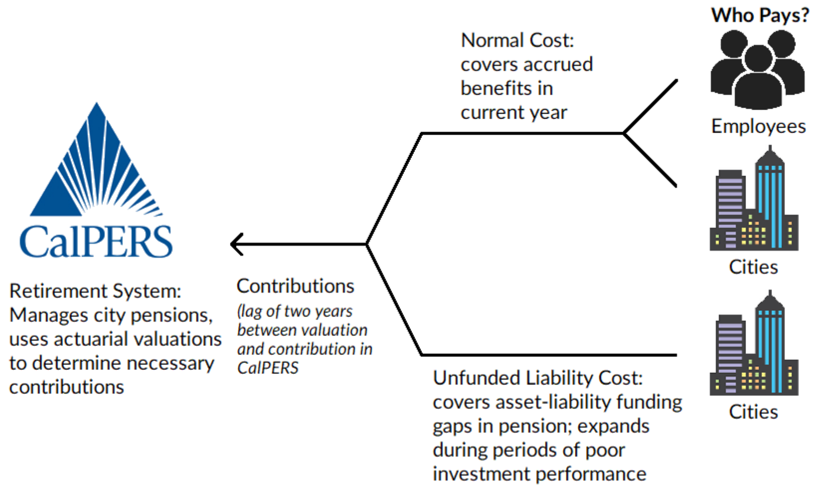


Figure 1.2: Municipal Pension Assets and Liabilities in California

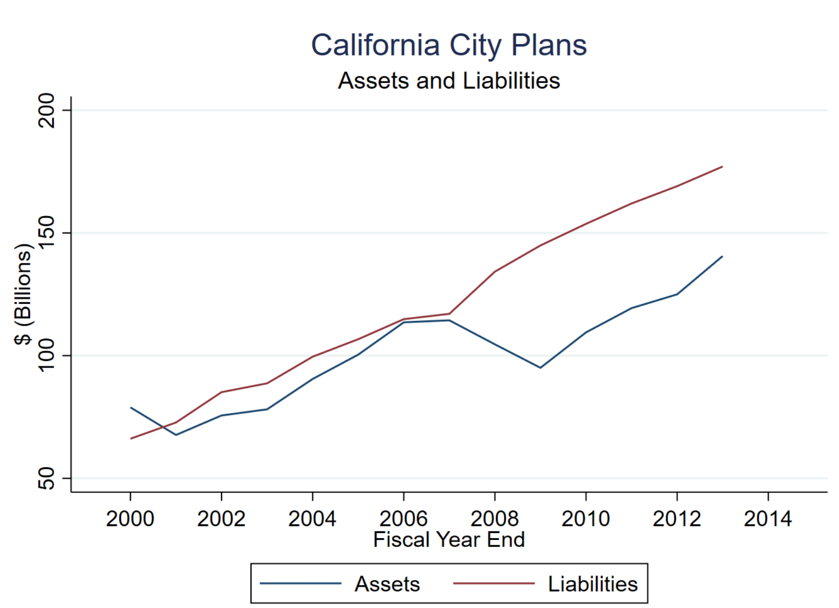
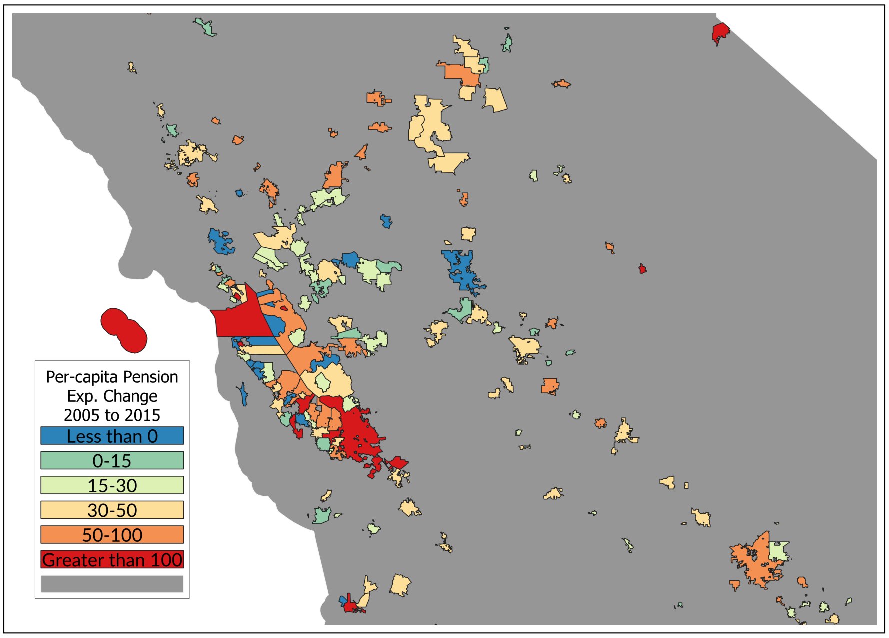
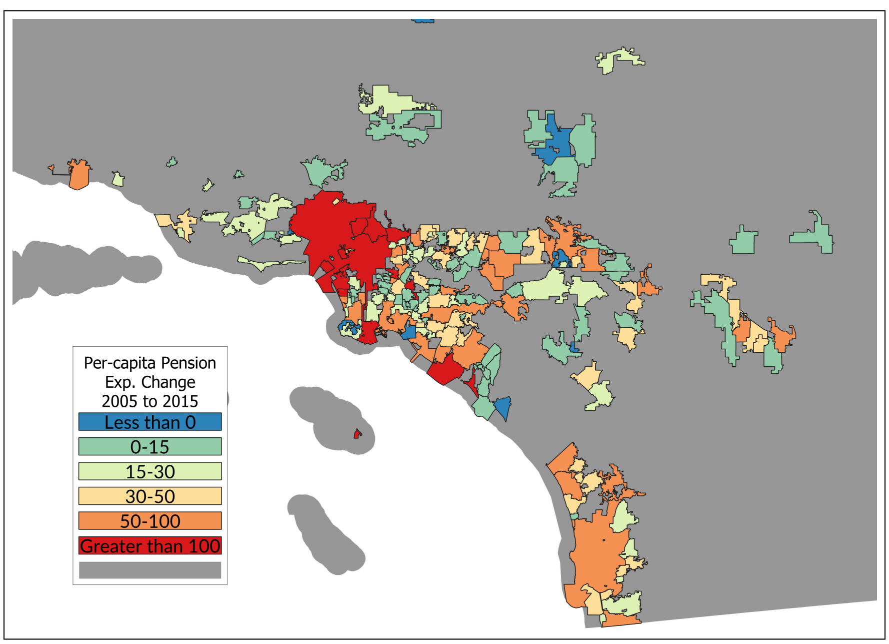


Figure 1.3: Central California: 2005-2015 Change in Pension Pressure



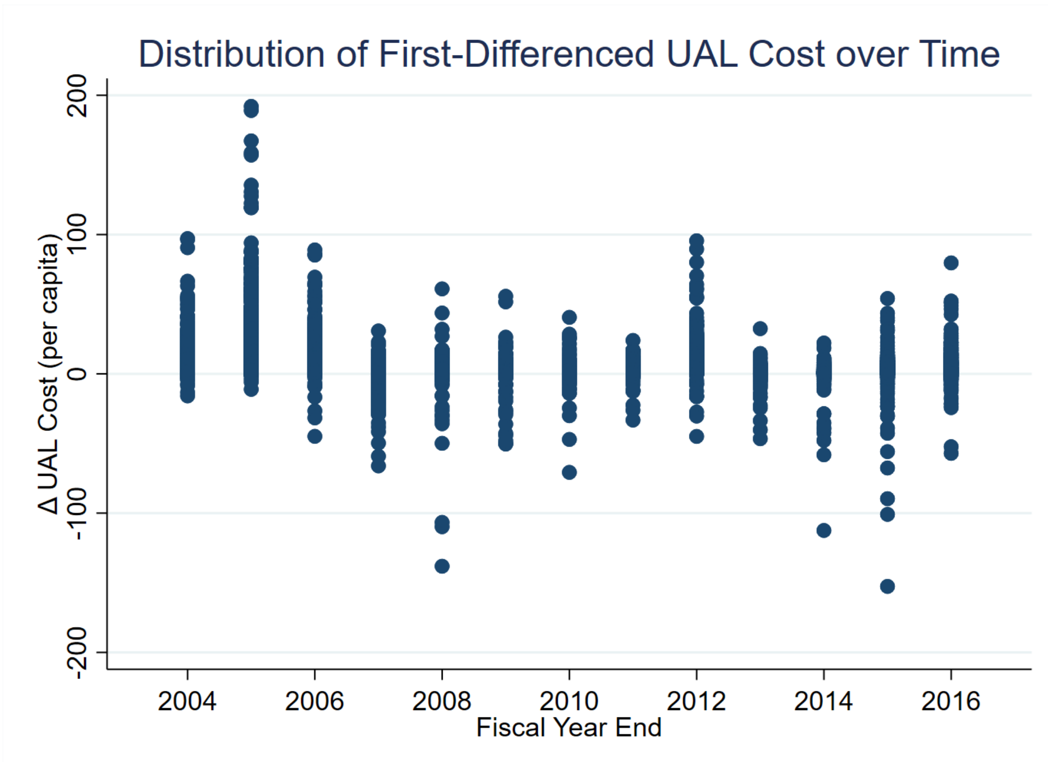
Source: Researcher-collected California pension data on CalPERS and other systems.

Figure 1.4: Southern California: 2005-2015 Change in Pension Pressure



Source: Researcher-collected California pension data on CalPERS and other systems.

Figure 1.5: Distribution of First-differenced UAL Cost over Time



Notes: "UAL Cost" refers to the payment a city is required to make to pay down its unfunded actuarial liability. Source: Researcher-collected California pension data on CalPERS and other systems.

1.8 Tables

Table 1: Definitions of Categorical Expenditures

Category	Definition
Public Safety	Expenditure to protect city residents. Includes law enforcement, fire suppression and prevention, and emergency medical services, along with various other services like animal regulation.
General Government	Expenditure to maintain functioning of the city government. Includes government officials (e.g., the city clerk) and their staff as well as administrative support services (e.g., budgeting and finances).
Transportation	Expenditure to facilitate the movement of people and goods. The primary components are streets, street landscaping and drainage, parking, and public transit.
Community Development	Expenditure to support the current and long-term economic wellbeing of the city. Includes planning, construction regulation, redevelopment, public housing, and community promotion.
Culture and Recreation	Expenditure to provide cultural and recreational opportunities. Category includes parks, libraries, public pools, museums, and community centers.
Health	Expenditure for sanitation and human health. Category includes sewers, solid waste removal, hospitals, and cemeteries.

Notes: Sourced from the Cities Financial Transactions Report (FTR) Instructions, which are given to cities by the California State Controller's Office to create uniform financial reports.

Table 2: Summary Statistics for City Expenses

	2005		2015	
	Per-capita	Share	Per-capita	Share
	(1)	(2)	(3)	(4)
Total	1648		1642	
	(1332)		(1313)	
Current	1299	0.809	1312	0.807
	(966)	(0.117)	(1113)	(0.119)
Non-current	349	0.191	329	0.193
	(502)	(0.117)	(372)	(0.119)
I. Current: Total				
Wages	412	0.271	411	0.249
	(358)	(0.101)	(382)	(0.086)
Benefits	185	0.113	216	0.129
	(162)	(0.056)	(216)	(0.055)
Normal Cost	44	0.026	47	0.029
	(39)	(0.014)	(42)	(0.013)
UAL Cost	11	0.006	45	0.027
	(21)	(0.011)	(50)	(0.018)
Private	216	0.141	257	0.153
	(272)	(0.112)	(406)	(0.113)
Other	477	0.296	446	0.286
	(446)	(0.137)	(409)	(0.140)
II. Current: Categorical				
Safety	433	0.286	442	0.297
	(319)	(0.108)	(323)	(0.106)
General Government	162	0.107	174	0.110
	(171)	(0.074)	(220)	(0.071)
Transportation	137	0.092	139	0.085
	(163)	(0.64)	(240)	(0.058)
Comm. Development	142	0.092	129	0.079
	(155)	(0.059)	(195)	(0.060)
Culture/Recreation	171	0.103	148	0.088
	(230)	(0.080)	(202)	(0.067)
Health	201	0.112	201	0.115
	(310)	(0.113)	(316)	(0.117)
III. Non-current				
Capital	270	0.152	229	0.136
	(430)	(0.112)	(293)	(0.107)
Debt	79	0.039	100	0.057
	(187)	(0.043)	(150)	(0.063)
N (cities)	412	412	428	428

Notes: Standard errors in parentheses. Expenditure variables are presented as per capita, 2016 dollars and as shares of total expenditures. Definitions of categorical expenses are provided in Table 1.

Table 3: Summary Statistics for Public Safety Employment and Crime

	2005		2015	
	Per 100k	Obs.	Per 100k	Obs.
	(1)	(2)	(3)	(4)
I. Public Safety Employment				
Total Police	210.8	291	183.6	302
	(120.3)		(123.2)	
Police, Officers	151.9	291	134.1	302
	(90.5)		(95.4)	
Police, Non-officers	58.9	291	49.5	302
	(40.1)		(36.4)	
Fire	95.6	215	94.9	201
	(66.8)		(77.8)	
EMS	34.2	122	46.8	114
	(50.2)		(68.2)	
II. UCR Crime and Arrest Rates				
Violent Crime	1054.4	408	862.9	413
	(628.3)		(556.2)	
Property Crime	3271.1	408	2406.6	413
	(1808.5)		(1403.0)	
Violent Arrests	601.2	408	535.0	413
	(444.4)		(407.0)	
Property Arrests	519.0	408	438.2	413
	(879.7)		(587.6)	

Notes: Standard errors in parentheses. Employment variables are rescaled per 100,000 residents of a city. Observations listed are cities for which police, fire, or EMS services, respectively, have employees which are paid directly by the city. Police employment data is sourced from the Law Enforcement Officers Killed in Action (LEOKA) files, while fire and EMS employment data are from the cities' Financial Transaction Reports. Crime and arrest counts for index offenses are from the UCR Return A files.

Table 4: Pension UAL Payment on Retirement and Benefit Spending (first differences)

	(1) Base	(2) Trended Covars	(3) City FE	(4) City, Year FEs	(5) Unweighted
	Panel A. Valuation UAL Cost on FTR Retirement Spending				
UAL Cost	0.873^{***}	0.811^{***}	0.814^{***}	0.760^{***}	0.837^{***}
	(0.086)	(0.090)	(0.096)	(0.108)	(0.203)
	[4807]	[4807]	[4805]	[4805]	[4807]
	Panel B. Persistence of UAL Cost on FTR Retirement Spending				
UAL Cost, t-1	0.166^{***}	0.045	-0.008	-0.134	0.146^{*}
	(0.062)	(0.074)	(0.076)	(0.122)	(0.082)
	[4432]	[4432]	[4430]	[4430]	[4432]
UAL Cost, t-2	0.029	-0.015	-0.059	0.014	0.145
	(0.054)	(0.057)	(0.061)	(0.118)	(0.172)
	[4055]	[4055]	[4050]	[4050]	[4055]
UAL Cost, t-3	0.062	0.090	0.079	-0.060	-0.057
	(0.058)	(0.064)	(0.067)	(0.086)	(0.225)
	[3673]	[3673]	[3671]	[3671]	[3673]
Model versions:					
Covariates	✓	✓	✓	✓	✓
Trended Covariates		✓	✓	✓	✓
City FE			✓	✓	
Year FE				✓	
Unweighted					✓

Notes: First-difference estimates of the impact of city unfunded actuarial liability (UAL) payment on city retirement spending, as reported in their Financial Transaction Reports. All regressions, except those indicated otherwise, are weighted by the city's average population during the sample period. All financial variables are in per capita, 2016 dollars. Standard errors in parentheses, clustered by city. Sample sizes are in brackets for each regression. The coefficients' significance levels are symbolized as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data cover 2003-2016 from California Financial Transaction Reports (FTR) and researcher collected pension data.

Table 5: Results: First-differences, Current and Noncurrent Expenditure

	(1) Base	(2) Trended Covars	(3) City FE	(4) City, Year FEs	(5) Unweighted
Total Expenditure	-0.178	-0.659	-0.722	-0.808	-0.641
	(0.388)	(0.442)	(0.472)	(0.675)	(0.609)
	[5281]	[5281]	[5280]	[5280]	[5281]
Deficit	0.071	-0.063	-0.039	1.002	-0.810
	(0.491)	(0.581)	(0.621)	(0.808)	(0.563)
	[5281]	[5281]	[5280]	[5280]	[5281]
Current	0.766***	0.436*	0.419	0.299	0.472*
	(0.243)	(0.234)	(0.255)	(0.390)	(0.254)
	[5281]	[5281]	[5280]	[5280]	[5281]
Wages	-0.212***	-0.343***	-0.343***	-0.327***	-0.261**
	(0.062)	(0.072)	(0.077)	(0.095)	(0.117)
	[5281]	[5281]	[5280]	[5280]	[5281]
Benefits	0.876***	0.773***	0.769***	0.733***	0.808***
	(0.099)	(0.104)	(0.105)	(0.136)	(0.145)
	[5281]	[5281]	[5280]	[5280]	[5281]
Retirement	0.873***	0.811***	0.814***	0.760***	0.837***
	(0.086)	(0.090)	(0.096)	(0.108)	(0.203)
	[4807]	[4807]	[4805]	[4805]	[4807]
Contracting	0.024	-0.001	-0.129	0.148	0.192
	(0.257)	(0.259)	(0.252)	(0.391)	(0.240)
	[4747]	[4747]	[4743]	[4743]	[4747]
Other	0.066	-0.002	0.114	-0.254	-0.299
	(0.410)	(0.398)	(0.407)	(0.571)	(0.263)
	[5281]	[5281]	[5280]	[5280]	[5281]
Non-current	-0.944***	-1.095***	-1.141***	-1.106**	-1.113
	(0.351)	(0.373)	(0.403)	(0.539)	(0.695)
	[5281]	[5281]	[5280]	[5280]	[5281]
Debt Service	-0.212	-0.261	-0.297*	-0.365**	-0.390
	(0.146)	(0.168)	(0.177)	(0.184)	(0.302)
	[5281]	[5281]	[5280]	[5280]	[5281]
Capital	-0.732**	-0.834***	-0.844**	-0.742	-0.723
	(0.305)	(0.314)	(0.337)	(0.472)	(0.495)
	[5281]	[5281]	[5280]	[5280]	[5281]
Model versions:					
Covariates	✓	✓	✓	✓	✓
Trended Covariates		✓	✓	✓	✓
City FE			✓	✓	
Year FE				✓	
Unweighted					✓

Notes: First-difference estimates of the impact of city unfunded actuarial liability (UAL) payment on city current and noncurrent expenditures. All regressions, except those indicated otherwise, are weighted by the city's average population during the sample period. All financial variables are in per capita, 2016 dollars. Standard errors in parentheses, clustered by city. Sample sizes are in brackets for each regression. The coefficients' significance levels are symbolized as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data cover 2003-2016 from California Financial Transaction Reports (FTR) and researcher collected pension data.

Table 6: Results: First-differences, Categorical Expenditure

	(1) Base	(2) Trended Covars	(3) City FE	(4) City, Year FEs	(5) Unweighted
Public Safety	0.483***	0.396***	0.381***	0.316**	0.194
	(0.087)	(0.097)	(0.101)	(0.136)	(0.131)
	[5281]	[5281]	[5280]	[5280]	[5281]
General Government	0.277**	0.313**	0.355***	0.348**	0.161
	(0.120)	(0.125)	(0.131)	(0.158)	(0.103)
	[5281]	[5281]	[5280]	[5280]	[5281]
Transportation	0.006	-0.012	-0.019	-0.055	0.274
	(0.052)	(0.063)	(0.075)	(0.103)	(0.204)
	[5281]	[5281]	[5280]	[5280]	[5281]
Community Development	0.061	0.036	0.009	0.067	0.090
	(0.112)	(0.126)	(0.130)	(0.221)	(0.082)
	[5281]	[5281]	[5280]	[5280]	[5281]
Culture/Recreation	0.212	0.225*	0.276**	0.182	-0.155
	(0.130)	(0.132)	(0.137)	(0.186)	(0.194)
	[5281]	[5281]	[5280]	[5280]	[5281]
Health	-0.188	-0.228	-0.247	-0.203	0.025
	(0.145)	(0.150)	(0.161)	(0.190)	(0.144)
	[5279]	[5279]	[5278]	[5278]	[5279]
Model versions:					
Covariates	✓	✓	✓	✓	✓
Trended Covariates		✓	✓	✓	✓
City FE			✓	✓	
Year FE				✓	
Unweighted					✓

Notes: First-difference estimates of the impact of city unfunded actuarial liability (UAL) payment on city outcomes. All regressions, except those indicated otherwise, are weighted by the city's average population during the sample period. All financial variables are in per capita, 2016 dollars. Standard errors in parentheses, clustered by city. Sample sizes are in brackets for each regression. The coefficients' significance levels are symbolized as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data cover 2003-2016 from California Financial Transaction Reports (FTR) and researcher collected pension data.

Table 7: Results: First-differences, Public Safety Employment

	(1) Base	(2) Trended Covars	(3) City FE	(4) City, Year FEs	(5) Unweighted
Police Emp. (LEOKA)	-0.045***	-0.068***	-0.062***	-0.051*	-0.054**
	(0.015)	(0.016)	(0.017)	(0.026)	(0.024)
	[3709]	[3709]	[3707]	[3707]	[3709]
Police Officers (LEOKA)	-0.032***	-0.046***	-0.044***	-0.040**	-0.040***
	(0.009)	(0.010)	(0.011)	(0.017)	(0.012)
	[3709]	[3709]	[3707]	[3707]	[3709]
Police Civilians (LEOKA)	-0.013	-0.022**	-0.018*	-0.011	-0.013
	(0.010)	(0.010)	(0.010)	(0.015)	(0.022)
	[3709]	[3709]	[3707]	[3707]	[3709]
Fire Emp.	0.009	-0.008	-0.003	-0.009	0.039
	(0.015)	(0.017)	(0.018)	(0.030)	(0.065)
	[2599]	[2599]	[2592]	[2592]	[2599]
EMS Emp.	-0.019	-0.024	-0.020	-0.009	0.003
	(0.023)	(0.024)	(0.026)	(0.030)	(0.022)
	[1309]	[1309]	[1288]	[1288]	[1309]
Model versions:					
Covariates	✓	✓	✓	✓	✓
Trended Covariates		✓	✓	✓	✓
City FE			✓	✓	
Year FE				✓	
Unweighted					✓

Notes: First-difference estimates of the impact of city unfunded actuarial liability (UAL) payment on city public safety employment outcomes. All regressions, except those indicated otherwise, are weighted by the city's average population during the sample period. All financial variables are in per capita, 2016 dollars. Employment outcomes are per 100,000 city residents. Standard errors in parentheses, clustered by city. Sample sizes are in brackets for each regression. The coefficients' significance levels are symbolized as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data cover 2003-2016 from California Financial Transaction Reports (FTR), researcher collected pension data, and Law Enforcement Officers Killed in Action (LEOKA) files.

Table 8: Results: First-differences, Crime Rates

	(1) Base	(2) Trended Covars	(3) City FE	(4) City, Year FEs	(5) Unweighted
Crime Cost	0.335**	0.455***	0.515***	0.346**	0.142
	(0.132)	(0.138)	(0.144)	(0.151)	(0.198)
	[5177]	[5177]	[5176]	[5176]	[5177]
Violent Crime Rate	0.345*	0.486***	0.562***	0.451**	0.071
	(0.179)	(0.188)	(0.195)	(0.202)	(0.288)
	[5177]	[5177]	[5176]	[5176]	[5177]
Property Crime Rate	2.483***	3.088***	3.301***	0.979	2.302***
	(0.486)	(0.500)	(0.532)	(0.679)	(0.487)
	[5177]	[5177]	[5176]	[5176]	[5177]
Violent Arrest Rate	-0.018	0.024	0.110	0.228	-0.097
	(0.176)	(0.180)	(0.187)	(0.240)	(0.226)
	[5177]	[5177]	[5176]	[5176]	[5177]
Property Arrest Rate	-0.142	-0.226	-0.197	-0.022	-0.175
	(0.163)	(0.175)	(0.179)	(0.240)	(0.292)
	[5177]	[5177]	[5176]	[5176]	[5177]
Violent Clearance Pct.	-0.016	-0.019*	-0.016	-0.001	-0.000
	(0.010)	(0.011)	(0.012)	(0.015)	(0.011)
	[5106]	[5106]	[5105]	[5105]	[5106]
Property Clearance Pct.	-0.017***	-0.023***	-0.024***	-0.006	-0.020***
	(0.004)	(0.005)	(0.005)	(0.007)	(0.006)
	[5107]	[5107]	[5106]	[5106]	[5107]
Model versions:					
Covariates	✓	✓	✓	✓	✓
Trended Covariates		✓	✓	✓	✓
City FE			✓	✓	
Year FE				✓	
Unweighted					✓

Notes: First-difference estimates of the impact of city unfunded actuarial liability (UAL) payment on city crime outcomes. All regressions, except those indicated otherwise, are weighted by the city's average population during the sample period. All financial variables are in per capita, 2016 dollars. Crime and arrest rates are per 100,000 residents. "Crime Cost" is per capita, and assumes \$67,794 and \$4,064 are the costs of the average violent and average property crimes, respectively. Standard errors in parentheses, clustered by city. Sample sizes are in brackets for each regression. The coefficients' significance levels are symbolized as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data cover 2003-2016 from researcher collected pension data and the Uniform Crime Reporting Return A files.

1.9 Additional Tables

Table 9: Results: First-differences, Debt

	(1) Base	(2) Trended Covars	(3) City FE	(4) Unweighted	(5) >100k Pop
Num. of Outstanding Debts	-0.000	-0.004	-0.002	-0.001	0.078**
	(0.006)	(0.007)	(0.007)	(0.003)	(0.034)
	[4744]	[4744]	[4741]	[4744]	[733]
Any Debt Issued	0.003*	0.003**	0.004**	0.003***	-0.002
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
	[5281]	[5281]	[5280]	[5281]	[735]
Amount Debt Issued	1.105**	1.150**	1.297**	0.811	-0.306
	(0.517)	(0.543)	(0.591)	(1.079)	(0.383)
	[5281]	[5281]	[5280]	[5281]	[735]
Outstanding Principal	-2.031*	-3.185***	-3.102***	-2.050	-4.987**
	(1.129)	(1.110)	(1.149)	(1.336)	(2.154)
	[4744]	[4744]	[4741]	[4744]	[733]
Model versions:					
Covariates	✓	✓	✓	✓	✓
Trended Covariates		✓	✓	✓	✓
City FE			✓		
Unweighted				✓	
Large cities					✓

Notes: First-difference estimates of the impact of city unfunded actuarial liability (UAL) payment on city outcomes. All regressions, except those indicated otherwise, are weighted by the city's average population during the sample period. All financial variables are in per capita, 2016 dollars. Standard errors in parentheses, clustered by city. The coefficients' significance levels are symbolized as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data cover 2003-2016 from California Financial Transaction Reports (FTR) and researcher collected pension data.

Table 10: Results: First-differences, Tax

	(1) Base	(2) Trended Covars	(3) City FE	(4) Unweighted	(5) ¿100k Pop
Num. of Tax Measures	-0.001	-0.001	-0.001	-0.001	0.004*
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
	[5281]	[5281]	[5280]	[5281]	[735]
Any Tax Measures	-0.002	-0.002	-0.002	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
	[5281]	[5281]	[5280]	[5281]	[735]
Model versions:					
Covariates	✓	✓	✓	✓	✓
Trended Covariates		✓	✓	✓	✓
City FE			✓		
Unweighted				✓	
Large cities					✓

Notes: First-difference estimates of the impact of city unfunded actuarial liability (UAL) payment on city outcomes. All regressions, except those indicated otherwise, are weighted by the city's average population during the sample period. All financial variables are in per capita, 2016 dollars. Standard errors in parentheses, clustered by city. The coefficients' significance levels are symbolized as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data cover 2003-2016 from California Financial Transaction Reports (FTR) and researcher collected pension data.

1.10 Comparisons to other City Data

To my knowledge, the California FTR data has not been used before. One reason may be that it has only recently come available; the change to the California Government Code which provides for the FTRs to be published digitally only came into effect after January 1, 2016. Without a demonstration of its substance, researchers may be unlikely to use it.

Researchers typically view the Census of Governments (CoG) and the associated Annual Survey of State and Local Government Finances as good sources of information on the behavior of cities. So, it provides a good dataset to compare the California FTRs to. Note that I cannot combine the CoG and Annual Surveys to get any sizeable panel of cities, since the Census only occurs every 5 years and the Annuals Surveys have relatively small and random samples. Still, I compare the CoG and FTR data where possible, to be described below.

The California State Controller provides a lengthy document instructing cities how

Table 11: Key Results: Robustness and Heterogeneity

	First Difference				Levels	
	(1) City FE	(2) Trimmed, 5%	(3) > Median Assets	(4) < Median Assets	(5) City, Year FEs	(6) With City Time Trends
Deficit	-0.039	0.111	-0.125	1.634	-0.557	-0.851
	(0.621)	(0.777)	(0.650)	(2.501)	(0.550)	(0.942)
	[5280]	[4962]	[2682]	[2598]	[5280]	[5280]
Current	0.419	0.136	0.364	1.075**	0.553	1.468*
	(0.255)	(0.304)	(0.273)	(0.540)	(0.650)	(0.806)
	[5280]	[4962]	[2682]	[2598]	[5280]	[5280]
Retirement	0.814***	0.788***	0.826***	0.610***	0.841***	0.931***
	(0.096)	(0.104)	(0.103)	(0.117)	(0.080)	(0.122)
	[4805]	[4532]	[2433]	[2372]	[4879]	[4879]
Wages	-0.343***	-0.429***	-0.338***	-0.377**	-0.266	-0.290**
	(0.077)	(0.085)	(0.081)	(0.177)	(0.217)	(0.128)
	[5280]	[4962]	[2682]	[2598]	[5280]	[5280]
Non-current	-1.141***	-1.016***	-1.211***	0.071	-0.512	-1.473**
	(0.403)	(0.382)	(0.425)	(0.904)	(0.582)	(0.631)
	[5280]	[4962]	[2682]	[2598]	[5280]	[5280]
Debt Service	-0.297*	-0.311	-0.326*	0.265	0.251*	-0.126
	(0.177)	(0.214)	(0.187)	(0.305)	(0.132)	(0.149)
	[5280]	[4962]	[2682]	[2598]	[5280]	[5280]
Capital	-0.844**	-0.704**	-0.885**	-0.193	-0.762	-1.347**
	(0.337)	(0.315)	(0.358)	(0.825)	(0.559)	(0.633)
	[5280]	[4962]	[2682]	[2598]	[5280]	[5280]
Police Emp. (LEOKA)	-0.062***	-0.050***	-0.061***	-0.066	-0.130***	-0.060**
	(0.017)	(0.012)	(0.018)	(0.058)	(0.035)	(0.025)
	[3707]	[3430]	[2497]	[1210]	[3731]	[3731]
Crime Cost	0.515***	0.584***	0.571***	-0.006	0.133	0.396*
	(0.144)	(0.173)	(0.151)	(0.476)	(0.258)	(0.204)
	[5176]	[4860]	[2658]	[2518]	[5177]	[5177]
Violent Crime Rate	0.562***	0.654***	0.658***	-0.459	0.111	0.523*
	(0.195)	(0.235)	(0.203)	(0.682)	(0.350)	(0.280)
	[5176]	[4860]	[2658]	[2518]	[5177]	[5177]
Property Crime Rate	3.301***	3.468***	3.072***	7.497***	1.424	1.014
	(0.532)	(0.606)	(0.551)	(1.423)	(0.931)	(0.701)
	[5176]	[4860]	[2658]	[2518]	[5177]	[5177]

Notes: Shows the impact of city unfunded actuarial liability (UAL) payment on city outcomes, where the specific outcomes in each row are key results drawn from earlier in the paper. Each column presents the results from either first-differenced regressions (columns 1-4) or level fixed-effect regressions (columns 5-6). Column 1 reproduces the results from previous tables from a model in first-differences with a city fixed-effect. Column 2 repeats this but trims the top 5% and bottom 5% of observations in terms of the change in the city UAL cost. Columns 3 and 4 restrict the sample to cities which held above and below median assets per capita, respectively, in 2003. Column 5 and 6 instead present results from level fixed-effect regressions, where column 5 has city and year fixed-effects and column 6 adds a city-specific linear time trend. All regressions, except those indicated otherwise, are weighted by the city's average population during the sample period. All financial variables are in per capita, 2016 dollars. Standard errors in parentheses, clustered by city. Sample sizes are in brackets for each regression. The coefficients' significance levels are symbolized as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data cover 2003-2016 from California Financial Transaction Reports (FTR) and researcher collected pension data.

to categorize budget items in the FTRs.[23] Similarly, the Census guides cities in answering the survey with the several hundred page Government Finance and Employment Classification Manual. Since cities differ in organizational and reporting structure, these survey instructions endorse comparisons closer to apples-to-apples than would otherwise be possible. Within a year and within a dataset, cities should be comparable, provided they respond accurately. However, the FTRs and CoGs have different categories and may use different definitions even for like categories. For instance, 'Parks and Recreation' in the CoG includes all "recreational and cultural-scientific facilities" like public parks, marinas, stadiums, and museums (but not libraries, which are separate). 'Parks and Recreation' in the FTRs is much more limited to maintenance of monuments and open spaces and expenses for athletics. With this in mind, I compare operating expenses for a few areas of interest - police, fire, health, parks and recreation, and libraries - using the 2012 Census of Governments and California FTRs from 2012. Out of these, police and fire are the most comparable. In the following table, I present summary statistics for the categories from each dataset as well as a percentage difference variable generated as

$$\%Diff_{c,2012} = \frac{|FTR_{c,2012} - CoG_{c,2012}|}{max\{FTR_{c,2012}, \quad CoG_{c,2012}\}}.$$

In general, the FTRs consistently report more cities than the CoG in 2012. Two categories, police and fire operating expenditures, have a high number of observations that are exactly the same, followed by libraries with around half of them matching. Further, for police over 90% of the values are within 5% of one another. The categories of 'Parks and Recreation' and 'Health' tend to have different values in the FTRs and the CoG for the reasons mentioned previously; the original 'Health' variable in the FTRs includes sewers and waste disposal, which is separate in the CoG. In my main dataset, I remove sewers and waste as well, since these tend to be operated as enterprises by the city.

[23]Cities Financial Transactions Report Instructions, California State Controller's Office.[24]https://www.sco.ca.gov/Files-ARD-Local/LocRep/Cities%20FTR%20Instructions.pdf

Table 12: Comparison of FTR and CoG City Categorical Expenses, 2012

Category	Data	Obs.	Mean	Std. Dev.	Min.	Max.	# Exact Matches
Police	FTR	481	20.2	95.9	0.015	1971.5	
	CoG	429	20.4	86.1	0.015	1664.7	
	% Diff.	429	0.041	0.155	0	0.948	361 of 429
Fire	FTR	326	12.0	38.4	0.001	556.1	
	CoG	307	11.7	31.7	0.001	422.5	
	% Diff.	300	0.016	0.102	0	0.996	287 of 300
Parks & Rec.	FTR	482	7.8	28.4	0	476.1	
	CoG	461	6.4	21.7	0.001	318.3	
	% Diff.	456	0.183	0.235	0	1	95 of 456
Libraries	FTR	188	4.7	14.4	0.001	137.9	
	CoG	183	3.6	9.8	0.002	95.1	
	% Diff.	178	0.100	0.211	0	0.994	86 of 178
Health	FTR	482	17.0	127.4	0	2553.1	
	CoG	356	4.8	46.6	0.001	847.6	
	% Diff.	356	0.815	0.271	0	1	6 of 356

All amounts are in 1,000,000s of dollars, barring the percent differences. Expenses are categorical operational expenses as they are defined in the Financial Transaction Reports (FTR) Instructions and Census of Governments (CoG) Government Finance and Employment Classification Manual for the fiscal year 2011-2012.

I also compared the cities with extreme differences in FTR and CoG values from each category to their respective Comprehensive Annual Financial Report (CAFR) values. For the CAFRs, I attempted to acquire them from the city's website, with mixed success. Some cities post the past decade of financial statements, while others post nothing. Since it would be a process to contact each city's officials to get them otherwise, building a dataset from CAFRs would be difficult. From tables where I compare the FTR, CoG, and CAFR data to one another, it looks like the FTRs better reflect the information presented in a city's audited financial statements (where they are available). For brevity, I only include one of these covering 'police'.

Table 13: Comparison of 2012 Police Expenditures, Extrema

City	FTR	CoG	CAFR
Lompoc	9.54	19.08	unavailable
Healdsburg	4.15	8.30	7.01[1]
Santa Cruz	19.78	39.56	unavailable
Highland	6.95	0.36	unavailable
San Clemente	12.10	0.65	11.98
Laguna Woods	1.30	0.11	1.62[1]
Rosemead	6.66	0.61	7.52[1]
Cerritos	13.56	1.31	unavailable
Montague	0.21	0.03	unavailable
Santa Fe Springs	8.80	1.61	unavailable
San Jacinto	8.62	1.64	unavailable
La Mirada	7.45	1.62	8.33[1]
Lancaster	26.40	6.10	23.49[1]
Imperial Beach	6.85	1.69	10.23[1]
Lakewood	10.88	2.92	12.28[1]

All amounts are in 1,000,000s of dollars. Comparison of Financial Transaction Reports (FTR) and Census of Governments (CG) data with Comprehensive Annual Financial Report (CAFR). This table looks at cities where the CG and FTR differ by a large margin in fiscal year 2011-2012.

Includes fire expenses with police as “Safety”.

1.11 Further Information on CalPERS Data

The California Public Employees' Retirement System (CalPERS) is the largest retirement system in the state of California, and services the majority of city's retirement plans, who contract with CalPERS for their benefits. In my study, it is therefore critical to know plan-level detail from these contracting city agencies; otherwise, I would be limited to a sample consisting of cities with independently serviced pensions, which tend to be in high population, established areas due to historical reasons related to the County Employees Retirement Law of 1937.

However, the readily available data on the Secretary of State's website was missing most city-plan-years from the years 2005-2016, and CalPERS does not immediately provide this information. To gather the necessary information, I submitted a public information request on CalPERS's online portal, asking for all records on each CalPERS contracting agency from 2003 to 2017 in order to match to my city financial data.[25] The records I obtained were "Annual Valuation Reports", which as per the language of the documents contain "important actuarial information about [the contracting agencies'] pension plan at CalPERS." These contain information about each plan's standing and expected future contributions. All in all, there were a little over 17,000 actuarial valuations, each representing a different plan-year observation from the 15 years of interest. Using Python, I turned these PDFs into text streams which I subsequently scraped into CSV files. The Annual Valuation Reports were variable both within and between years in terms of their length and contents. To help give a sense of what these reports look like, I present two pages from the Annual Valuation Report for Riverside, CA for 2005 below.

More specifically, I deemed that there are two general templates for the reports: 'short' and 'long'. 'Short' reports were provided to plans which in 2003 had less than 100 members, whereas the 'long' reports were provided to plans with 100 or more members. This

[25]CalPERS: Public Records Requests https://www.calpers.ca.gov/page/contact/public-records-requests

separation stemmed from CalPERS pooling these smaller employer-plans together in order to reduce the size of fluctuations in their year-to-year contributions. While possibly a boon by providing more predictable budgeting in smaller agencies, risk pooling is a bane for a researcher trying to build a data set. From 2004-2010, 'short' valuations do not report assets and liabilities on a per-plan basis. Instead, there is only the value of a "Side Fund", which is the difference in fundedness between risk pool and plan, as well as plan membership and payroll. Separately in Risk Pool Valuations, information on the risk pool's financial status is provided in aggregate each of these years. In 2011, they again calculate assets and liabilities for all plans, using the following formula to reestablish plan fundedness

$$\mathbf{MVA_{plan}} = \frac{\mathbf{Liabilities_{plan}} + SideFund_{plan}}{Liabilities_{pool} + SideFund_{pool}} * MVA_{pool}.$$

For me to calculate the fundedness for the period 2004-2010, I need either of the two bolded quantities in the equation above so that I can get the other. Since I have data at either end (for the majority of plans), I choose to interpolate the liabilities between the two points, and then calculate the MVA_{plan} using the equation above. This assumes liabilities move linearly, which in aggregate is mostly the case. Assets are allowed to move up and down with macroeconomic shocks through the MVA_{pool} term, and incorporating plan Side Funds means plans vary in fundedness appropriately. The end result is having information for the CalPERS contracting agencies for about ˜96% of the plan years after interpolation, and about three-quarters before.

Chapter 2

Politically Charged: District Attorney Partisanship, Dismissal Rates, and Recidivism

2.1 Introduction

Each year, over 17 million criminal cases enter the United States court system—at least one case for every 15 American adults (Court Statistics Project 2018). The sheer volume of cases empowers local district attorneys (DAs) with near-total discretion to choose how many cases on which they want to pursue conviction, at the expense of scarce judicial resources, and how many to dismiss, at a potential social cost if non-conviction spurs future crime (Alschuler 1968; Bibas 2004; Bowers 2010; Stith 2008).[1] The decision to pursue a conviction carries significant consequences for accused individuals, who face economic and social fallout from incarceration and accruing a criminal record (Agan and Starr 2018; Dobbie, Goldin, and Yang 2018; Kling 2006; Mueller-Smith 2015).

[1] As we note in Section 2, in our sampled jurisdictions, as in most criminal justice systems in the United States, the vast majority of filed criminal cases either result in dismissal—whereby the accused person is released without penalty—or conviction. Vanishingly few cases result in a not-guilty verdict.

But a DA's choice of conviction rate also has political ramifications. Elected prosecutors must answer to voters who, polls indicate, consistently see high crime rates as a major problem (Gramlich 2016; McCarthy 2020).[2] A median voter framework suggests that all DAs, regardless of their own political identities, might pursue high conviction rates as a visible signal of their toughness to the median, crime-averse voter. Conversely, as "citizen-candidates" (Besley and Coate 1997), DAs might incorporate their own partisan perceptions of crime rates and deterrence, topics on which Democrats' and Republicans' views diverge (Gramlich 2021; McCarthy 2020; Yokley 2021). Surprisingly, given this theoretical ambiguity and DAs' central role in the criminal justice system, there exists no rigorous evidence showing how elected prosecutors' political identities affect their decision over whether to seek conviction or dismiss a case.

This paper examines the causal impact of DA partisan affiliation on case dismissal rates, sentencing outcomes, and re-offense rates. Our goal is to identify differences in how Democratic and Republican DAs exercise their prosecutorial authority, and ultimately assess how those differences shape the efficacy of their local justice systems. Prior work suggests that DA partisanship influences local incarceration patterns: Arora (2019) and Krumholz (2020) show that electing Democratic DAs leads to fewer prison admissions. However, these studies lack the criminal case-level data necessary to observe dismissal rates or re-offense behavior. And while recent anecdotal evidence suggests that "progressive prosecutors" pursue fewer cases (following campaign promises to do so), these examples capture only a subset of Democratic DAs drawn mostly from large urban jurisdictions where Democratic voters hold substantial majorities.[3] Both of these confounding factors could independently influence conviction rates. Thus, it remains an open question whether DA partisanship matters, both at the margin and on average.

[2]By "district attorney," we mean, broadly, the chief prosecuting authority in a given jurisdiction; some states refer to these officials by other names, such as "state's attorney" or "commonwealth attorney." The majority of district attorneys nationwide are elected in partisan contests, although some states elect public prosecutors through nonpartisan ballots.

[3]Agan, Doleac, and Harvey (2021b) inventory many of these reform-minded prosecutors.

Our study bridges these empirical gaps. We deploy a multi-state administrative dataset that contains over ten million individual criminal case records spanning 2000-2019.[4] Crucially, these data include, to the best of our knowledge, data on all criminal cases filed in our six sampled states, including defendants whose cases were ultimately dropped, allowing us to accurately measure rates of case dismissal. Matching these court data to elections returns, we can also observe the partisanship of the serving DA at the time of case filing.

However, jurisdictions that elect Democratic DAs differ systematically from those that elect Republican DAs—particularly when they win office in uncontested elections, as is the case for 81 percent of the elected DAs in our sample. Specifically, we find that jurisdictions served by Democratic prosecutors have noticeably higher caseloads (both in absolute terms and relative to their population) than those served by Republican DAs. The net bias of that added caseload is unclear: jurisdictions that handle more cases likely have more resources to do so, while higher caseloads could put pressure on DAs to quickly dismiss more low-level offenses in order to focus on relatively serious cases.

To address this endogeneity concern, we use narrowly-decided DA elections to home in on otherwise comparable jurisdictions that barely elect DAs from each party. Logically, by focusing on the most politically competitive jurisdictions where both parties have a shot of winning elections, we minimize the baseline differences between court systems in our sample served by Democratic and Republican DAs.[5] Indeed, we find that winning DA partisanship in narrowly-decided elections is uncorrelated with election-year jurisdiction and case characteristics, which suggests that any post-election differences in case outcomes represent the causal impact of DA partisanship. Our preferred research design combines this close-election, regression discontinuity-style (RD) variation with a difference-in-differences

[4]Specifically, we use data from Arkansas, Colorado, Kentucky, North Carolina, Texas, and Virginia, states for which we could obtain comprehensive criminal justice data and DA election outcomes. While these states are mostly in the South, they include roughly one-quarter of all DAs nationwide, as well as a range of urban and rural jurisdictions that in many ways resemble the country more broadly. We detail our sample of court systems and elections in Section 2.

[5]See Dippel (2022) and Macartney and Singleton (2017) for recent examples of studies that follow this close-elections approach to estimate the policy impact of electing partisan officials.

specification, which isolates the effect of DA partisanship both within jurisdictions across time and across jurisdictions with different election outcomes. Additionally, we supplement our main close-elections approach with a matching design, leveraging our detailed case-level data to examine the external validity our close-election identification strategy. That is, we compare outcomes for similar cases from similar jurisdictions that are prosecuted by DAs from opposing parties, irrespective of their election margins, to consider the average impact of DA partisanship.

Our close-elections difference-in-differences estimates show that, relative to Republican DAs, the marginal Democratic DA is 25 percent (8 percentage points) *more* likely to dismiss incoming criminal cases. Similarly, cases filed in jurisdictions led by Democratic DAs are 17 percent (9.7 percentage points) less likely to result in incarceration. We also find that electing a Democratic DA leads to 32 percent shorter incarceration sentences on average (about 1 less month spent in jail or prison), in line with prior research.[6] These effects remain consistent under different choices of controls and definitions of "close" elections; heterogeneity analyses point to larger effects among nonwhite and female defendants, as well as among those accused of felony and nonviolent offenses, although these differences are not statistically significant. Moreover, our matching estimators back up our main findings. Within our primary close-elections research sample, our matching approach produces comparable results to our preferred specification, which we argue attests to its validity. Expanding our attention to cases from all jurisdictions, matching estimates indicate that the average Democratic DA dismisses 18 percent more cases—a smaller impact than the marginal Democratic DA, but economically meaningful nonetheless. We interpret these results as evidence that Democratic DAs pursue higher dismissal rates and impose less punitive sentences across a range of jurisdictions and defendants.

[6]Prosecutors do not unilaterally impose sentences, a power that rests with judges. Still, by choosing whether and which charges to pursue, DAs effectively determine the minimum and maximum jail or prison sentences that a judge could choose from. DAs can also recommend particular punishments as part of a plea deal. Our point estimate on sentencing lies almost exactly between those of Arora (2019)—who finds that Democratic DAs generate roughly 55 percent shorter incarceration sentences—and Krumholz (2020), who finds that Democratic DAs reduce total sentenced months in a jurisdiction by 6 percent.

Democratic DAs' high dismissal rates clearly serve accused individuals' interests, avoiding the social and economic penalties that come with a criminal record. Yet, it remains unclear whether this agenda serves the interest of justice. Recent studies suggest that, at the margin of dismissal, a prosecutor's decision to drop a case could actually promote public safety, in that non-prosecution reduces the probability that arrestees commit future crimes. In particular, Agan, Doleac, and Harvey (2021a) show that dismissing misdemeanor charges reduces the probability of re-arrest, while Augustine et al. (2022) and Mueller-Smith and Schnepel (2019) find that diverting felony defendants diminishes the likelihood that they face future criminal charges.[7] Motivated by these findings, we examine how Democratic DAs affect the rate at which affected individuals recidivate and reappear in future criminal cases, as a way of gauging DAs' impact on the efficacy of the local court system.

We find that the marginal Democratic DA has no statistically or economically significant effect on the probability of re-offense, measured within 1 or 2 years of initial case filing, relative to the marginal Republican DA. Point estimates have negative signs and fairly tight confidence intervals, such that we can rule out increases in recidivism of more than 2 percent, both within 1 year and 2 years of initial arrest (or 0.6 and 0.8 percentage-point increases, respectively).[8] We stress that these estimates do not capture the causal effect of case dismissal per se; rather, our results just capture the reduced-form effect of Democratic district attorneys on recidivism rates—which could operate through numerous policy channels, including higher case dismissal rates. Still, our null findings echo an emerging empirical consensus that criminal convictions might have limited deterrence effects at the margin.

Overall, our work sheds new light on how partisan politics impacts the criminal justice system. This analysis speaks to two strands of the literature. First, by providing rigorous evidence that a district attorney political identities substantially influence case dismissal

[7] As we discuss in Section 2, we consider diversion—typically probation, followed by charge dismissal—to be a form of case dismissal.

[8] Certain specifications yield precisely-estimated declines in recidivism of up to 16 percent (6 percentage points) among cases filed four years after a Democratic DA wins a close election. We treat these results with caution, as we discuss in the text, although they broadly support our argument that Democratic DAs do not adversely affect re-offense rates.

rates, we contribute to a growing body of research that highlights the degree to which case outcomes in the legal system depend not on the facts but on idiosyncratic courtroom actors, including judges (Cohen and Yang 2019), defense attorneys (Agan, Freedman, and Owens 2021; Shem-Tov 2022), and assistant prosecutors (Sloan 2020; Tuttle 2019). Second, we provide a new perspective on the long-standing question of whether a less punitive criminal justice system comes at the cost of reduced deterrence. Across a range of jurisdictions, we show how Democratic DAs pursue fewer criminal convictions and lower incarceration rates without systematically adverse effects on recidivism. While we cannot directly estimate the impact of case dismissal on re-offense behavior, our results support the argument—advanced by academics and policymakers alike—that high conviction rates are not necessarily efficient. Altogether, our findings underscore the extent to which politics shapes the implementation of the law at the local level, and how the punitiveness of the court system hinges on the partisanship of district attorneys.

2.2 Background and Data

Our study examines the causal link between DA partisan affiliation and the rate at which they dismiss criminal cases. To capture results that can be plausibly attributed to DA partisanship, and to deliver widely-applicable findings, we compile a multi-state dataset that links DA election returns with detailed criminal justice records. However, few states provide the comprehensive criminal justice records necessary to conduct an analysis focused on pre-sentencing outcomes, a critical constraint which limits the scale of this study. As we detail below, our final dataset includes six states with suitable policy environments and data—Arkansas, Colorado, Kentucky, North Carolina, Texas, and Virginia.[9] Together, these states comprise 17.8 percent of the U.S. population as of the 2010 Census and around a

[9] As of July 19, 2023, the findings presented in this paper do not include Kentucky, in order to comply with requirements from the Administrative Office of the Courts. Including Kentucky—which is a small part of our sample anyway—has virtually no effect on our results. We will include these findings in the final version of this paper.

quarter of its district attorneys.[10] Though we attempted to collect criminal justice data for the period 2000-2018, many states do not have records from the early 2000s. As such, data availability varies across states: Arkansas records span 2000-2018, Colorado 2002-2018, Kentucky 2002-2018, North Carolina 2013-2018, Texas 2000-2019, and Virginia 2008-2018.

We begin by contextualizing district attorneys' role in the criminal process. We then summarize the DA elections data that we use to identify the partisanship of DAs serving in our sampled jurisdiction-years before describing how we construct our sample of criminal court records. Combining these datasets, we summarize key features of our sampled criminal cases and jurisdictions, highlighting differences between Democrat- and Republican-led court systems. Finally, we discuss how we construct our primary analytic dataset, a jurisdiction-election panel.

2.2.1 Background: The Role of District Attorneys in the Criminal Justice Process

Voters in the six states we study elect district attorneys to lead public prosecutors' offices and represent the state against defendants in felony and most misdemeanor cases. Each district attorney serves as the chief law enforcement officer in their judicial district, usually a single county, but sometimes a grouping of small counties.[11] Many jurisdictions staff public prosecutors' offices with assistant district attorneys (ADAs), who often function as the actual prosecuting attorneys on cases. Even in these settings, though, the district attorney retains control of overall prosecutorial policy and directs the ADAs. In our sampled states, voters elect DAs to office every four to six years.

District attorneys' principal responsibility involves deciding whether to seek to convict

[10] We also collected criminal justice data from North Dakota, Oregon, and Pennsylvania, but could not use them in this study, due to quality issues (Pennsylvania) and lack of data to identify DA partisanship (North Dakota and Oregon). Our resulting dataset overlaps extensively with those in recent research that, like us, attempt to gather court records from as many states as possible (for example, Dippel and Poyker [2019] and Feigenberg and Miller [2021]).

[11] For example, Texas Judicial District 97 in the northern part of the state includes the counties of Archer, Clay, and Montague, which have a collective population of around 38,000.

individuals accused of a criminal offense. Following an arrest by the police, DAs can opt to pursue the arresting charge, impose alternative or further charges on an individual based on the evidence, or dismiss the case against the person altogether.[12] DAs face few constraints on their prosecutorial discretion. In fact, recent reformist prosecutors have declined to prosecute whole categories of offenses (e.g., drug possession). Should they decide to pursue a defendant's conviction, DAs also have leverage over the eventual sentence imposed, both via the choice of specific charge to prosecute and through the plea bargaining process, which resolves the vast majority of cases. Their wide remit positions DAs to substantially influence criminal justice outcomes in their local court systems.

2.2.2 Data: District Attorney Elections

For each jurisdiction and year represented in our sampled states, we determine the political party of the serving DA using election records. We compile jurisdiction-level DA election returns, which include candidate-level vote totals, and use the partisan identity of the most recent election winner as the party of the sitting DA. For five of the six states in our sample, we directly observe candidates' party affiliations; for one state (Arkansas), we determine candidates' political parties using voter registration data.[13] From this information, we calculate the winning party's margin of victory as a share of the total votes cast in the election.[14] Note that one state in our sample (Kentucky) reports only competitive election results, though, as we discuss in Section 4, the omission of non-competitive elections does not affect our primary research design.

[12] In five of our six sampled states, DAs formally charge arrestees. The exception is Virginia, where the police decide whether to charge. Still, in all states, DAs have the authority to dismiss a case entirely, with no further penalty.

[13] Arkansas is the only state in our sample that holds nonpartisan DA contests but has partisan voter registration. As such, we match candidates to the Arkansas voter roll to determine their registered political party preference. We find that many contested elections feature candidates from opposing parties—however, our ability to observe candidate partisanship hinges on the match quality to the Arkansas voter roll, which may introduce bias. Omitting Arkansas from our sample has a negligible impact on our findings.

[14] For elections with more than two candidates, we define the margin of victory as the difference between the first- and second-placed candidates' vote shares. We separately define the winning *party's* margin of victory as the difference between the top Democratic candidate's vote share and the top Republican's vote share.

In Table 1, we summarize all 1,311 DA elections that took place in our sampled states and time periods. Republican candidates won most (55 percent) of these contests, while Democrats won 43 percent of the time. In the remaining 2 percent of elections, either a third-party candidate won, or else we could not determine the partisanship of the winning candidate. Strikingly, only 255 races (18 percent of the total) were contested by multiple candidates.[15] And even among these nominally competitive elections, most races had wide victory margins: on average, the winning candidate defeated her opponent by 19 percentage points. Focusing on the 197 "politically competitive" races that include at least one Democrat and one Republican—the sample of greatest interest to this study—we find that the average race resulted in a Democratic loss of 6 percentage points. In Section 3, we discuss in greater detail how we leverage these election margins to identify the causal effect of electing a Democratic DA.

2.2.3 Data: Criminal Justice Records

Our primary research question considers the rate at which DAs dismiss criminal cases. Many criminal justice datasets only include criminal offenses that result in conviction and sentencing. We therefore compile charge-level administrative records that describe all felony and misdemeanor charges filed in our sampled states. These data contain, to the best of our knowledge, all criminal charges filed in each state during their coverage periods, including those that district attorneys ultimately chose to dismiss. That said, arrest and charging behavior varies substantially across states. Unequal arrest and charging rates on the part of the police mean that identical crimes committed in different states may not be reported at similar frequencies. Likewise, the same underlying offense can often support different charges in different states. We quantify and discuss the ramifications of these cross-state differences below.

[15]Krumholz (2019) collects data on DA elections held in 40 states between 1990 to 2015, finding that 25 percent of races were contested, slightly higher than the rate in our sample, but confirmation that DA races rarely feature multiple candidates.

Criminal Justice Data Cleaning

Each state data source records fairly detailed information about individual charges, such as a description of the offense, its severity (felony, misdemeanor, or infraction), the date the charge was filed, the date the charge was disposed, the charge disposition (e.g., dismissed, found or pleaded guilty), sentencing outcomes associated with the charge, and characteristics of the defendant (age, gender, race/ethnicity). Our approach to data cleaning echoes Feigenberg and Miller (2021), who compile a similar multi-state dataset.[16]

We first define charge-level characteristics and outcomes. Following the Universal Crime Reporting (UCR) system's offense classification scheme, we label charges as either property, violent, drug, or traffic offenses. We refer to charges that do not fall into these categories—generally, crimes against society, such as driving while intoxicated (DWI)—as "other." For each charge, we define three disposition outcomes: an indicator for whether the charge was dropped or dismissed by the DA; an indicator for whether the charge resulted in any incarceration sentence; and the nominal length of that sentence, in months.[17,18]

We then aggregate our charge data to the criminal-case level, our primary unit of obser-

[16]While our overall approach to data cleaning resembles that of Feigenberg and Miller (2021), two differences stand out. First, to define criminal cases, we group together charges based not only on identifying fields in the data but also on date of filing, a choice which we believe reflects how the courts will process these charges. This choice reduces the number of "cases" in our sample, relative to that of Feigenberg and Miller. Second, likely as a result of how we define cases, we observe incarceration rates 8-10 percentage points higher than Feigenberg and Miller, and dismissal rates that are 6-8 percentage points lower. We believe both approaches to the data are reasonable ex ante; we prefer our approach because, by defining cases relatively broadly, we avoid mistaking a case dismissal for charges dropped as part of an overall guilty plea on a case (anecdotal and empirical evidence suggests that such "charge bargaining" is ubiquitous in the criminal justice system; see, for example, Piehl and Bushway [2007]).

[17]Different court systems refer to non-prosecution dispositions by different names, which we group under the term "dismissed." We consider a charge to be dismissed if its disposition is given as "dropped," "dismissed," "*nolle prosequi*," or, specifically in Texas, "deferred adjudication." "Deferred adjudication" is a diversion outcome unique to felony offenses filed in that state, which amounts to probation followed by dismissal, provided the defendant does not re-offend. Note that we exclude from this definition the rare situations in which charges were dropped because a judge dismissed a case (for example, due to a lack of evidence), rather than a DA.

[18]We do not observe the actual sentences served by defendants, but rather the incarceration term set at the time of sentencing. External factors unobservable to us, such as parole board leniency, could cause nominal and actual incarceration sentences to diverge in ways that local DAs (as well as judges and defendants) may anticipate and factor into their decision-making. For this reason, we tend to stress our extensive-margin results, which consider dismissal and the probability of any incarceration sentence.

vation, using defendant identifiers and charge filing dates. We assume that the courts will process charges filed on the same day for the same defendant together as a single unit. We categorize cases as property, violent, drug, traffic, or other if any charge within that case is of the given type. Likewise, we define a felony case to include at least one felony charge.

To construct our primary case-level outcomes—case dismissal, incarceration, and sentence length—we again look across charges within a case. We consider a case to be dismissed if all charges on the case are dismissed; we say a case results in incarceration if at least one charge results in incarceration; and we take the incarceration length imposed on a case to be the maximum confinement sentence imposed across all charges (which we set to zero if no charge on the case resulted in an incarceration term). Since our sentence length variable has a large variance and a mode of zero, we prefer to use its inverse hyperbolic sine (asinh), which has a similar interpretation and properties as the natural log function, but is well-defined at zero.[19]

We then construct defendant covariates, relying on fields in the raw data to identify defendant gender and age at the time of case filing.[20] Each state reports defendant race/ethnicity identification differently: some treat "Hispanic" as a race, mutually exclusive with "white," while others treat ethnicity as distinct from race. To minimize measurement error, we group together "nonwhite" defendants, which includes defendants identified as Hispanic or by any race other than "white."[21]

Finally, we use our extensive data and defendant identifiers to track re-offense behavior. For each defendant-case (i.e., every appearance a defendant makes in the data), we create indicators for whether the defendant re-appears on a new criminal case within 1 or 2 years their original case's filing date, which we refer to as defendant recidivism. Unfortunately,

[19] Feigenberg and Miller (2021) and Shem-Tov (2021) use this transformation as well to measure length of incarceration. We obtain virtually identical results when we instead use a typical log transformation.

[20] Court records from Virginia provide no age or year or birth field, and so we omit this variable for Virginia defendants.

[21] Functionally, this approach mirrors that of Feigenberg and Miller (2021), who face a similar challenge and focus on a white/nonwhite dichotomy. Unlike those authors, though, we include less-represented races in our sample, such as Asian and American Indian.

only three states (Kentucky, North Carolina, and Texas) provide sufficient data to track defendants across time and observe future offenses. Still, these three states account for over 60 percent of the criminal cases and elections in our main research sample.

To ensure that our dataset contains comparable types of cases across jurisdictions, we impose several sample restrictions. We drop cases that have indeterminate dispositions—such as cases binding to a higher court—or no disposition. We omit cases that exclusively involve traffic offenses, since some states do not report low-level traffic infractions and misdemeanors. We also exclude juveniles, those younger than 18 at the time of case filing. When considering defendant recidivism, we exclude cases filed within 1 or 2 years (depending on the exact outcome variable) of the end of their state's criminal justice records' coverage period.

Summary of Criminal Case Data

Our final sample contains 9,787,182 criminal cases, which we summarize in Table 2. Across our full dataset (described in column 1), DAs dismiss 37 percent of cases, while 47 percent of cases result in incarceration. Though not shown here, virtually all cases that do not result in dismissal culminate in conviction; as such, we typically refer to defendants whose cases were not dismissed as "convicted." In terms of recidivism, 41 percent of defendants re-appear on a subsequent criminal case within two years. Defendants are overwhelmingly male (just 25 percent are identified as female), while 54 percent are identified as nonwhite. Most cases involve misdemeanors (only 34 percent include a felony charge), and the most common charge types included in cases are property offenses (e.g., theft) and "other" offenses against society (e.g., DWI), which appear in 36 and 39 percent of cases, respectively.

The remaining columns of Table 2 underscore the extent to which defendant outcomes diverge across court systems. Breaking down our sample by state, we find that case dismissal rates range from 19 percent (Arkansas) to 58 percent (North Carolina); incarceration rates range from 15 percent (North Carolina) to 62 percent (Texas). These stark disparities across state lines likely reflect a combination of factors. As we mention above, despite the fact that our data comprise arrest-to-sentencing records, standards for arrest, and what sorts of charges police and prosecutors have at their disposal, vary across jurisdictions. Equally, actual judicial punitiveness almost certainly varies across court systems, driven by racial attitudes (as Feigenberg and Miller [2021] demonstrate), mandatory sentencing laws, and—of particular relevance for us—the political considerations of judges and district attorneys.

Regardless of its underlying causes, the immense cross-state variation in criminal case composition and outcomes poses an empirical challenge as we try to disentangle the impact of DA partisanship from other confounding factors. In general, our preferred approach is to focus on within-jurisdiction variation, usually in the form of jurisdiction fixed effects, that account for unrelated state-by-state (and, indeed, jurisdiction-by-jurisdiction) differences in the legal environment. We elaborate on how our research design isolates the effect of DA partisanship in Section 3.

2.2.4 Constructing Panel Datasets

We combine our DA election and criminal justice records to create two analytic datasets. The first uses the elections returns to create a jurisdiction-level panel, which tracks the partisanship of each jurisdiction's serving DA over time, as well as basic information from the most recent DA election (e.g., the winning candidate's margin of victory). Using county and year of filing, we match individual cases to this panel of jurisdictions, which allows us to observe the DA's partisanship at the time a case entered the court system.

We summarize our jurisdiction-year panel in Table 3, highlighting both demographic features of the locations represented in our data, as well as the average characteristics of the local court systems. To help illustrate the (qualitative) relationship between DA partisanship and features of their jurisdictions, we break down our sample of jurisdiction-years by the party affiliation of the serving DA, as well as the competitiveness of the election in which they secured their office. The first panel of Table 3 summarizes jurisdiction-year demographics, drawing from Census (population, nonwhite population share) and Bureau of Economic

Analysis (income per capita) data.[22] Column 1 describes all jurisdiction-years in our sample (N=3,874); the remaining columns separate jurisdiction-years by the partisanship of the sitting DA, as well as by the competitiveness of the most recent DA election.

The average jurisdiction-year in our sample has a population of just over 171,000—24 percent of whom are nonwhite—and a mean per-capita income of almost $39,000. These averages mask substantial disparities across jurisdictions served by Democratic and Republican DAs: Democratic DAs hold office in less populous and less affluent places than their Republican peers. Interestingly, differences between contested and uncontested jurisdictions, regardless of DA partisanship, stand out, whereas contested jurisdictions represented by DAs of opposing parties appear relatively similar. This pattern presages our research design, which focuses on these competitive jurisdiction-elections.

The second panel of Table 3 focuses on the local court outcomes of systems represented by Democratic and Republican DAs. These data show further evidence of divisions between politically competitive and noncompetitive localities. Uncontested jurisdictions have relatively low caseloads (with fewer than 2,000 cases filed per year, compared around 5,000 in contested jurisdictions) and relatively high incarceration rates (40-48 percent, versus 39-45 percent in contested jurisdictions). Interestingly, Republican- and Democratic-led jurisdictions also appear to diverge in their punitiveness: contested jurisdictions with Democratic DAs have higher dismissal rates (39 percent versus 35 percent) and lower incarceration rates (39 percent versus 45 percent) than their Republican counterparts elected in competitive races.[23] At the same time, the average Democratic DA in a contested jurisdiction handles 6 percent more cases per year than the average competitively-elected Republican (5,267 cases per year, compared to 4,987 in Republican-led DAs offices), which encapsulates the fact

[22]Note that BEA data is missing for incorporated cities in Virginia, which operate independently from their surrounding counties and elect their own DAs. For jurisdictions that span multiple counties, we aggregate Census and BEA data from these different counties. We inflate BEA income data to 2016 dollars.

[23]We summarize other case and defendant characteristics by DA partisanship in the Appendix. The case-level summary data show similar patterns in judicial outcomes, with cases filed in Republican-led jurisdictions more likely to result in prosecution and incarceration. By contrast, we find few differences in average case characteristics.

that DAs of opposing parties serve very different court systems. The question thus remains whether this divergence in outcomes represent the causal impact of DA partisanship, or result from the underlying differences in caseloads and demographics across jurisdictions. Our identification strategy, which we discuss in the next section, aims to disentangle these factors from the causal impact of DA partisanship.

As part of that research design, we construct a second analytic dataset: a jurisdiction-election panel, focusing exclusively on contested DA races. Intuitively, as we discuss in depth in Section 3, by restricting our attention to these politically competitive locations, we hope to minimize differences in jurisdiction characteristics that might explain variation in court outcomes. Around each of the 197 "politically competitive" elections in our sample that feature both Democratic and Republican candidates, we create a panel that includes all criminal cases filed between the third year prior to the election and the sixth year following the election. Each observation in the panel is a case-by-jurisdiction-election. The panel has an unbalanced structure, where periods may be missing for some elections but not others. Within a given jurisdiction-election's panel, each criminal-case observation shares time-invariant, election-level data, such as the election margin. The dataset has a "stacked" structure, in which cases can appear more than once in the panels of different elections. This jurisdiction-election panel contains 8,196,244 case-by-jurisdiction-election observations.[24]

2.3 Research Design

Our goal is to estimate the causal effect of district attorney partisanship on the probability of criminal case dismissal, sentence length, and the likelihood of defendant re-arrest. Using a straightforward ordinary least-squares (OLS) framework, we could regress defendant outcomes on an indicator for whether a Democratic DA held office at the time of case filing.

[24] We describe post-election cases from this sample of contested jurisdiction-elections, alongside our full sample (discussed above) and our final research sample, in the Appendix.

However, that approach would yield a biased estimate if there is any correlation between the underlying determinants of judicial system punitiveness and the political identity of the local prosecutor. For example, Table 3 shows that, on average, jurisdictions with Democratic district attorneys elected in competitive races handle larger volumes of cases per capita than jurisdictions with Republican prosecutors elected in uncontested races. The direction of the resulting bias is unclear: those additional cases could reflect greater judicial capacity to pursue cases to conviction, or the high caseload could oblige DAs to dismiss more cases in order to conserve judicial resources for the most serious offenses. Disentangling the causal effect of DA partisanship from the role of prosecutorial resources and other confounding factors represents the primary goal of our research design.

2.3.1 Identification from Close Elections

To address this endogeneity concern, we focus on jurisdictions in which Democratic and Republican DAs hold office as the result of closely-contested elections. As Table 3 shows, jurisdictions served by Democratic and Republican DAs elected in contested races appear to be fairly similar, at least along observable dimensions. Our close-elections research design extends this insight to its logical conclusion: by focusing on those DAs elected in the most narrowly-decided races, we hope to minimize the confounding differences between jurisdictions served by Democratic and Republican chief prosecutors. This approach echoes numerous prior studies that use close elections to infer the causal effect of partisan officials.[25]

Cross-sectional Specification

Our most parsimonious empirical approach captures the effect of close Democratic victories in DA elections using a sharp regression discontinuity (RD) framework. For this specification, we aggregate our data to the jurisdiction-year level, and use the result of the most recent DA election as an instrument for the political identity of the serving DA at the

[25] See, for example, Dippel (2022), Ferreira and Gyourko (2009), Lee et al. (2004), and Macartney and Singleton (2019).

time of case filing.[26] For jurisdiction j and year t, we regress average case outcome $\overline{Y}$ (for instance, the share of cases dismissed by the DA) on an indicator for whether a Democratic candidate won the last DA election in j ($Democrat$), along with a linear control for the margin by which she won (or lost) that election, and their interaction.[27] Our main cross-sectional specification is

$$\overline{Y}_{jt} = \alpha_0 + \alpha_1 Margin_{jt} + Democrat_{jt}(\alpha_2 + \alpha_3 Margin_{jt}) + \epsilon_{jt}, \tag{2.1}$$

where the coefficient of interest, α_2, captures the local average effect of barely electing a Democratic DA on aggregate outcome $\overline{Y}$.

Throughout our analysis, we restrict our sample to those jurisdiction-elections decided by a sufficiently narrow margin to be comparable. Specifically, we focus on elections in which the Democratic candidate's margin of victory (or loss) was at most 7.1 percentage points (that is, those for which $Margin$ lies in the range [-0.071,0.071]). We use that threshold since it is the smallest—and therfore, we argue, most conservative—bandwidth given by Calonico, Cattaneo, and Farrell's (2020) selection procedure across our different samples and outcome variables.[28] Still, to address concerns that this sample restriction explains our findings, in Section 4 we show that we obtain similar results using narrower and wider definitions of "close" races. Across all specifications, we cluster our standard errors at the

[26]Our estimator is "intent-to-treat," insofar as we do not confirm the identity of the district attorney who actually prosecuted a given case. For instance, if a Democratic DA is voted into office, but is recalled after one year and replaced by an appointed successor, cases prosecuted in her jurisdiction will still be in our "treated" group. More generally, some cases filed around the time of an election will remain in the system when a new DA assumes office. Since we use the date of case filing to assign cases to DAs—rather than the date of case disposition, which is potentially endogenous (i.e., more case dismissals will shorten average disposition times)—these election-year cases could be misclassified as "treated" or "untreated." Since we observe that DA partisanship takes time to express itself on case outcomes, we do not believe the resulting bias substantially affects our results.

[27]To demonstrate the robustness of our results, we estimate an alternative specification that includes a quadratic polynomial in *Margin*. As we discuss in the next section, our results are qualitatively similar under both the linear and polynomial approaches.

[28]In the Appendix, we compare cases filed in all contested jurisdictions to those filed in our close-elections sample. We find few differences in average case characteristics, suggesting our sample of cases filed in narrowly-contested jurisdictions is observably similar to the larger sample of cases filed in competitive jurisdictions.

jurisdiction-by-election level to correct for the correlation of outcomes among cases handled by the same DA.[29] Our identifying assumption is that, among narrowly-decided contests, and conditional on the election margin, DA partisanship (*Democrat*) is uncorrelated with unobserved determinants of average case outcomes, ϵ.

Internal Validity of Close-elections Approach

While not testable in itself, our identification assumption implies that, pre-election, we should not find any observable differences distinguishing the jurisdictions where Democratic DA candidates barely won and lost. That is, close Democratic DA victories should not be correlated with features of the jurisdiction or its typical caseload that might independently drive post-election case outcomes. Using Equation 2.1, we estimate differences in election-year jurisdiction demographics, as well as case and defendant characteristics. We disaggregate case and defendant data here, since our primary research design (which we discuss below) leverages case-level data, although we obtain quantitatively similar results when we aggregate to the jurisdiction-election level. We report our estimates in Table 4. Encouragingly, we find no significant differences in these jurisdiction- or case-level covariates that would undermine our identification approach.

In keeping with standard RD assumptions, we also ensure that our running variable—the difference in vote share between the Democratic and Republican DA candidates in the election—is balanced at the cutoff separating Democratic and Republican victories. In other words, within our sample of 197 politically competitive elections, there should not be a discontinuity in the running variable density at the cutoff that would indicate one party systematically wins these tight races. Figure 2.2 shows that the distribution of our running variable (the Democratic margin of victory) does not vary discontinuously at zero. A formal test for any jump in the density at the cutoff (following Calonico, Cattaneo, and Farrell [2020]) fails to detect any such difference, with a p-value of 0.987.

[29]Functionally, clustering at the jurisdiction-election level amounts to clustering on the running variable ($Margin$).

External Validity of Close-elections Approach

Our close-elections comparison approach, combined with our policy setting, raises a potential empirical concern: we rely on a small group of jurisdictions for identification, with just 49 DA elections in our sample decided by a margin that falls within our bandwidth. As with all election-RD designs, our coefficient of interest recovers treatment effects among cases in marginal jurisdictions where candidates compete in close races. And, more specific to our context, the fact that most DA elections are not even contested could limit the external validity of our findings, as the effects we observe in competitive jurisdictions may not provide much insight into the numerous uncontested jurisdictions in our data. To help address these points, in Section 5 we introduce a supplementary matching design that, while perhaps offering a less airtight causal claim, allows us to comment on the average effect of DA partisanship across a wider range of jurisdictions.

2.3.2 Panel Specification

While intuitive, the cross-sectional approach given by Equation 2.1 might not yield very persuasive results, given the small sample of competitive elections at our disposal. Equally, cross-sectional estimates across jurisdictions must contend with the substantial variation in outcomes we see across court systems (see Table 2), which could lead to noisy results.

In fact, the traditional sharp RD approach leaves information on the table that could make our results both more precise and more compelling. Our panel structure allows us to compare defendant outcomes both across jurisdictions with different election outcomes as well as within jurisdictions across time, which can assist with statistical inference. Likewise, our case-level data allow us to control for a wide array of defendant and case characteristics to help discount alternative explanations for our results.

To strengthen our research design and leverage these features of our data, we combine our close-elections, RD-style intuition with a difference-in-differences panel framework. We estimate this panel model using our main disaggregated, case-level panel dataset, including pre- and post-election observations, and focusing on outcome Y for individual case i (for example, an indicator for whether the case was dropped). Building on our cross-sectional specification from Equation 2.1, we add controls for whether case i was filed in a post-election period τ ($post_{i\tau}$), as well as jurisdiction and year fixed effects (λ_j and θ_t, respectively). Our baseline panel specification is:

$$\begin{aligned} Y_{ijt\tau} =& \alpha_0 + \alpha_1 Margin_{jt} + Democrat_{jt}(\alpha_2 + \alpha_3 Margin_{jt}) + \\ & post_{i\tau}\Big(\gamma_1 + \gamma_2 Margin_{jt} + Democrat_{jt}(\beta_1 + \beta_2 Margin_{jt})\Big) + \\ & \lambda_j + \theta_t + \epsilon_{ijt\tau}, \end{aligned} \tag{2.2}$$

where β_1—the post-election effect of a Democrat winning in a close race—is the coefficient of interest.[30]

Intuitively, Equation 2.2 measures the difference-in-differences effect of a Democratic DA victory, relative to a Republican win, where the Democratic victory "treatment" is quasi-randomly assigned via close elections.[31] Including pre-election observations improves statistical efficiency, and addresses the possibility that pre-election circumstances drive our results. Furthermore, the addition of jurisdiction and year fixed effects help mitigate the substantial variation in outcomes across court systems and time (see Table 2), which further improves the precision of our estimates. For these reasons, we favor these case-level panel estimates. Still, we appeal to our aggregate cross-sectional results for transparency and to

[30]The estimated β_1 from Equation 2.2 without the panel components is effectively the same as α_1 from Equation 2.1 with jurisdiction-caseload weights. We prefer the case-level approach because it allows us to leverage our detailed data to achieve the greatest precision, as well as to thoroughly explore heterogeneity and alternative explanations for our findings. In any case, we obtain comparable estimates using both our cross-sectional and panel RD approaches, implying that our case-level design—and its associated weighting choice—does not drive our results.

[31]Similar hybridized RD-difference-in-differences approaches are increasingly common in research that examines the impact of local government politics and policies. Beach and Jones (2017) and Grembi, Nannicini, and Troiano (2016) employ similar "difference-in-discontinuities" designs, while Fischer (2023) and Shi and Singleton (2022) use analogous instrumental variables specifications with difference-in-differences components.

demonstrate the robustness our findings to alternative approaches.[32]

Finally, to more explicitly rule out pre-election trends in our outcome variables that might bias our results, and to comment on the dynamics of Democratic DAs' impact on criminal justice outcomes, we modify Equation 2.2 to take an event-study approach. This specification highlights period-specific impacts of DA partisanship, using fixed effects, $\kappa_{i\tau}$, to denote whether case i was filed in pre- or post-election period τ:

$$\begin{aligned} Y_{ijt\tau} =& \alpha_0 + \alpha_1 Margin_{jt} + Democrat_{jt}(\alpha_2 + \alpha_3 Margin_{jt}) + \\ & \sum_{\tau=-3}^{\tau=6} \Big(\kappa_{i\tau} + \rho_\tau Margin_{jt} + Democrat_{jt}(\delta_{1\tau} + \delta_{2\tau} Margin_{jt})\Big) + \\ & \lambda_j + \theta_t + \epsilon_{ijt\tau}. \end{aligned} \tag{2.3}$$

Each event-study coefficient $\delta_{1\tau}$ captures the difference in outcome Y between the last pre-election year (period -1) and period τ attributable to a marginal Democratic election victory.

2.3.3 Variation in DA Partisanship After Close Elections

Before estimating our cross-sectional and panel models, we first illustrate the "first stage" variation in district attorney partisanship that our close-elections design generates. In the cross section, our close-elections approach guarantees that, in the year after a DA election, jurisdictions in which a Democratic candidate won the race will be served by a Democratic DA. Over time, though, subsequent elections could result in incumbent losses, in both "treated" and "control" jurisdictions, which could narrow that gap.

To illustrate this pattern, we estimate an event study specification in the style of Equation 2.3, where the outcome is whether a Democratic DA holds office, and each observation is

[32] Qualitatively, we find similar results using alternative versions of Equation 2.2 that omit the panel framework. We provide these results in the Appendix to highlight the different roles played by our panel approach and choices of fixed effects. This comparison underscores the important role played by jurisdiction and year fixed effects in our empirical approach, particularly once we incorporate multiple post-election periods. As expected, our panel design—which includes pre-election observations—does not have a large effect on the magnitudes of our estimates, but does noticeably improve precision.

a jurisdiction-year. Those estimates appear in Figure 2.1. As expected, we find that during the four years after an election in which a Democratic candidate edged out their opponent, most jurisdictions continue to be served by a Democratic DA; after four years—at which point the modal jurisdiction holds another DA election—we find no significant differences in DA partisanship across locations that originally elected Democratic and Republican prosecutors. We interpret this pattern as evidence that our close-elections design succeeds in generating quasi-random differences in DA partisanship for a meaningful stretch of time. We note, too, that point estimates for the first four post-election years are virtually equal to 1, and there are no discernible pre-election trends, which affirms that the partisan identity of the pre-election DA is balanced across treated and untreated locations.

2.4 How Do Democratic DAs Affect Case Outcomes?

Our study investigates how DA partisanship causally influences criminal case outcomes. We begin our analysis by using our close-election framework to investigate how Democratic DAs affect the probability of case dismissal before broadening our attention to other salient case outcomes, namely the probability of incarceration and incarceration length. Anecdotal evidence surrounding "progressive prosecutors," plus our descriptive findings in Table 3, suggest Democratic DAs increase case dismissal rates while lowering incarceration rates and sentence lengths. Yet, as we have noted, specific features of Democrat-led jurisdictions—their lower income levels and higher caseloads, for instance—might explain these trends.

2.4.1 Does DA Partisanship Matter for Dismissal Rates? Visual Evidence

Using our close-elections framework and panel data, we first provide visual evidence of how narrow Democratic DA elections shape case dismissal rates in their jurisdictions.

Democratic DAs' election victories might take time to materially affect case outcomes—incoming DAs could, for example, need to replace assistant prosecutors, or lay out specific prosecution policies, two challenges faced by recent reform-minded DAs. We therefore begin by assessing the dynamic impact of close Democratic DA elections with an RD event study-style framework, using Equation 2.3 to estimate the treatment effect on case dismissal rates over time.

These estimates appear in the first panel of Figure 2.3. Importantly, we do not find any evidence of pre-election differences in case dismissal rates that would bias our results. The lack of differential pre-trends helps alleviate concerns that dismissal rates in treated and untreated jurisdictions might have already been diverging prior to electing a Democratic DA. Note that some case observations from the election year (period 0) will be "treated," in that they will still be pending when the new DA term begins in the subsequent year, which explains why we find a small (albeit insignificant) uptick in dismissal rates in the election year itself.[33]

Following the election, panel 1 of Figure 2.3 shows a rise in dismissal rates almost immediately after a Democratic DA assumes office, although the effects are only statistically significant beginning three years after the election. Four years after the election—at the conclusion of the modal DA's term in office—a marginally elected Democratic DA is 13 percentage points more likely to dismiss a criminal case than a marginally elected Republican DA.[34] Though noisy, point estimates indicate that this effect may attenuate 5-6 years after the election, which corresponds to the start of a new DA term in most jurisdictions.

In the second panel of Figure 2.3, we focus on that fourth post-election year, when the effects of DA partisanship appear most pronounced. We visualize the cross-sectional discontinuity in aggregate case dismissal rates that underpins our panel estimates, using a bin scatterplot to show the average dismissal rate across jurisdiction-elections by the Democratic

[33]For context, the average case in our sample takes 110 days to reach a disposition, meaning that the average case filed after mid-September of an election year will be disposed by a new DA, should the incumbent lose their re-election bid.

[34]Point estimates from this and other event study-style plots in the main text appear in the Appendix.

DA candidate's margin of victory (or loss).[35] This nonparametric approach reveals the stark variation behind our findings: at the cutoff, cases prosecuted by Democrats who just won their elections appear to have about a 12 percentage-point higher probability of being dismissed than those prosecuted by Republicans who just won their elections.

2.4.2 Estimating the Effect of DA Partisanship on Case Dismissal Rates

We next provide formal regression estimates to substantiate the visual evidence from Figure 2.3. Our estimates come from our two RD specifications: a typical cross-sectional RD estimator, following Equation 2.1, and an RD panel estimator, following Equation 2.2. We present these findings in the first panel of Table 5.

Results from our preferred panel RD specification appear in columns 2-4. These coefficients capture the difference-in-differences impact of a close Democratic DA's election over the 4 or 6 years post-election. Point estimates indicate that the marginal Democratic DA increases the probability of case dismissal by 7.4-8.3 percentage points during this period (26-28 percent of the sample mean), relative to the marginal Republican DA. Our estimates are robust to the inclusion of additional defendant and case covariates (column 3), as well as to different choices of post-election sample periods (compare columns 3 and 4).

In columns 5-7 of of Table 5, we present results from alternative cross-sectional specifications in order to demonstrate the robustness of our panel design, using Equation 2.1 and aggregate, jurisdiction-election-level data. To facilitate comparisons across jurisdictions that vary widely in size, we employ jurisdiction population weights (see Section 2.3 and the Appendix).[36] We first estimate a straightforward sharp RD model, examining the impact of

[35]We provide versions of all our bin scatterplots without the bins—that is, scatterplots where each point corresponds to exactly one jurisdiction-election—in the Appendix. These include figures in which the dependent variable is the within-jurisdiction change in mean outcomes, which we use in Table 5 below. These Appendix figures more directly mimic our cross-sectional RD approach by weighting observations by jurisdiction population.

[36]We use population as our weighting variable since it is unlikely to be influenced by the election outcome. Our panel regression design, which effectively weights jurisdiction-elections by caseload, arrives at comparable

a marginal Democratic DA victory in the fourth year post-election as in Figure 2.3 (N=49 jurisdiction-elections). That estimate appears in column 5. We find that the marginal Democratic DA causes an uptick in jurisdiction dismissal rates of 14.7 percentage points. Still, this effect is not precisely estimated— unsurprisingly, given the lack of controls for unrelated cross-state and over-time variation that our panel design includes.[37]

To mitigate these unrelated sources of variation, we present results from two alternative cross-sectional specifications. The first, shown in column 6 of Table 5, examines how a Democratic DA affects the *change* in dismissal rates between the pre-election and post-election periods within their jurisdiction (that is, the difference in average dismissal rates between the three years prior to the election and the four years after the election). This approach echoes the panel design from the second column of Table 5, but employs aggregate data (N=49 jurisdiction-elections) and no additional controls. We find that jurisdictions that barely elect a Democratic DA see a marginally significant 8 percentage-point increase in dismissal rates post-election. The estimate, which is quite similar to the baseline panel RD result from column 2, provides some reassurance that using case-level data and a formal difference-in-differences framework does not drive our results.

Lastly, to demonstrate that our findings do not hinge on the inclusion of pre-election data, we disaggregate our outcome variable slightly so that each observation is an election-period (N=282). This step means our observations are no longer unique at the jurisdiction-election level, and so we can augment our model with jurisdiction and year fixed effects, which control for secular variation in outcomes driven by differences in the legal environments across jurisdictions. The point estimate from that model, reported in column 7, indicates that Democratic DAs raise dismissal rates by 10 percentage points (31 percent)—slightly larger than the analogous panel estimate from column 4. Overall, we argue our results

estimates, suggesting our choice of weight does not drive our findings.

[37]Using bias-corrected standard errors, as prescribed in Calonico et al. (2020), does not substantially alter our conclusions from our cross-sectional RD results. Estimates from our baseline model (column 5) remain imprecise, while those in column 6 are actually more precise after using corrected standard errors. Note that we cannot use the same bias-correcting procedure with our panel RD specification, nor with the fixed effects specification shown in column 7.

provide robust evidence that DA partisanship meaningfully affects case dismissal rates, with Democratic DAs pursuing fewer convictions than their Republican counterparts.

2.4.3 Does DA Partisanship Matter for Other Case Outcomes?

Beyond case dismissal, we want to understand how DA partisanship shapes case dispositions broadly, particularly in terms of salient incarceration outcomes. Figure 2.4 presents visual evidence that electing a Democratic DA leads to lower rates of incarceration as well as lower incarceration sentences. Drawing on Equation 2.3, event study-style estimates in the first two panels point to marked declines in incarceration severity soon after a Democratic DA's election victory, culminating in a 14 percentage-point decline in incarceration probability and a 30 percent decline in sentence length in the fourth post-election year. Cross-sectional data in the third and fourth panels tell a similar story. However, the discontinuity in sentence length does not appear to be very robust near the cutoff, a point we return to below.

In the second panel of Table 5, we present panel and cross-sectional RD estimates describing Democratic DAs' impact on the probability of incarceration and incarceration sentence length. Using our main panel specification, we find that Democratic DAs reduce the probability of incarceration by between 9.2 and 9.8 percentage points in the 4-6 years following their election (16-17 percent of the sample mean). Likewise, we find that Democratic DAs impose 32-38 percent shorter sentences (roughly 3 fewer months in jail or prison, on average). Our cross-sectional RD approach, focusing on the fourth post-election year, shows an imprecise though economically meaningful decline in the probability of incarceration (around 8 percentage points), but near-zero effects on incarceration length. The imprecise null on sentence length likely reflects the noise near in the cutoff visible in Figure 2.4. When we consider the change in jurisdiction-wide dismissal rates, or include multiple post-election periods alongside jurisdiction and year fixed effects, we recover statistically significant Democratic DA effects on both incarceration probability and length. We conclude that DA partisanship plays a meaningful role in determining incarceration outcomes.

2.4.4 Robustness

Our twin panel and cross-sectional specifications provide some assurance that our choices of specification and observation level do not qualitatively affect our principal findings. That said, the RD framework underlying our panel approach requires assumptions over the exact sample and functional form of our control for the election margin, each of which could influence our findings. In Table 6, we further explore the robustness of our main close-elections panel estimates to different choices of sample and RD controls.

The second and third columns of Table 6 probe the robustness of our findings to alternative samples of elections. That is, we modify our bandwidth—our definition of what constitutes a "close election"—to include more or fewer jurisdiction-elections. In column 2, we narrow our sample to jurisdiction-elections decided by 5.7 percentage points or less (20 percent less than our primary bandwidth of 7.1 percentage points; this leaves us with 40 elections); in column 3, we widen our bandwidth to include jurisdiction-elections decided by 14.2 percentage points or less (twice our primary bandwidth; this gives us a total of 92 elections).[38] In general, our point estimates remain fairly stable and precise across these different samples. In column 4, we alter our panel RD specification (Equation 2.2) to include a quadratic, in addition to a linear, control for the Democratic margin of victory. This change has almost no effect on our results. Finally, to ensure our estimates do not depend on observations right at the cutoff—that is, DA elections decided by the slimmest margins, whose outcomes might be most susceptible to nonrandom factors—we exclude DA elections decided by 1 percentage point or less (that is, we take out a 1 percentage-point "donut"). We find uniformly larger treatment effects once we exclude the closest DA contests in our sample, which confirms our findings do not rely on those races.

[38]As hinted at earlier, in practice the "optimal" bandwidths for our different outcomes lie between 5.4 percentage points and 9.8 percentage points. Our choices of "narrow" and "wide" bandwidths here are meant to be salient but also extremes on that scale, to demonstrate that the choice of bandwidth does not materially affect our conclusions.

2.4.5 Heterogeneity

To round out our main findings, we examine whether Democratic DAs' choices of low prosecution and incarceration rates benefit some defendants or types of cases more than others. Documenting any heterogeneous effects of DA partisanship—or the lack thereof—provides useful context for interpreting our results: do Democratic DAs simply decline to prosecute specific types of cases (e.g., drug offenses), which single-handedly drives down prosecution rates, or does DA partisanship matter for a wide range of defendants?

We consider multiple dimensions of potentially heterogeneous treatment effects in Table 8. Each column presents panel RD results, applying our preferred model (Equation 2.2), estimated on a particular subsample of our data, described in the column headers. For context, below each point estimate and standard error we report the subsample mean of the outcome variable, which we find varies considerably.

Qualitatively, our heterogeneity analysis reveals larger effects of close Democratic DA elections on dismissal and incarceration rates among nonwhite and female defendants, as well as among cases involving felony and nonviolent offenses. We find larger impacts on sentence length among young and male defendants, as well as among cases involving felony and property offenses. In relative terms compared to their respective means, however, these effect sizes do not vary much, and there is no clear pattern to indicate that DA partisanship primarily impacts more or less severe cases, or more or less vulnerable defendants.

To test whether these differences in the treatment effect estimates across subsamples are statistically meaningful, we run a series of pooled specifications. That is, we modify Equation 2.2 to interact our treatment indicator (whether a Democrat wins the close DA election) with an indicator for the respective subsample (e.g., whether the defendant is over age 30). Broadly, we find little evidence of statistically significant differences. So while Table 8 points to larger effects among some types of cases and defendants, we cannot rule out equality in all but a few of these estimates. Based on these tests and the fairly consistent estimates in Table 8, our results suggest that Democratic DAs have a substantial effect across their caseload.

2.5 Do Democratic DAs Increase Recidivism Rates?

The differences in case outcomes generated by Republican and Democratic DAs that we have documented thus far point to substantially different official views on the value of criminal prosecution and incarceration. Ample prior research has shown that defendants incur considerable economic and social penalties following spells in detention, however brief, including wage penalties and lower civic engagement (Mueller-Smith 2016; White 2019). Although we do not have adequate data to examine those downstream effects of DA partisanship, we can still weigh in on the relative efficacy of Democratic DAs' policy choices by exploring how prosecutor partisanship influences re-offense rates.

A popular narrative suggests that, by prosecuting and potentially imprisoning defendants, DAs can deter future criminal behavior. A growing body of research has undercut this argument (for example, Agan et al. [2021a], Augustine et al. [2022], and Mueller-Smith and Schnepel [2020]). Nonetheless, these studies focus on the marginal non-prosecuted defendant arrested in specific, oftentimes coastal, jurisdictions. One might argue that a different, negative relationship between non-prosecution and recidivism holds in a the majority of smaller court systems, like those in our sample. We stress that we cannot directly test the impact of case dismissal on the probability that a defendant re-appears on a future criminal case. However, we can evaluate the effect of electing a Democratic DA—who, among other things, increases dismissal rates and lowers incarceration rates—on the overall re-arrest rate in their jurisdiction. This exercise helps provide both a sense of whether DA partisanship affects the deterrence of local criminal justice systems, as well as a descriptive data point on the relationship between relatively lenient prosecutors policies and re-offense rates.[39]

[39] We define recidivism as synonymous with re-arrest or, more precisely, re-appearance on a new criminal case. Many researchers oppose this definition, since arrest probability depends on the behavior of police. In

We apply our close-elections empirical framework to examine the impact of Democratic election victories on recidivism rates. We have two outcomes of interest, the probabilities a defendant re-offends within 1 year or 2 years of their original case's filing date.[40] Recall that we can only observe defendant re-arrests in three states, Kentucky, North Carolina, and Texas.

Figure 2.5 depicts the relationship between Democratic DA election victories and defendant recidivism, again using our (parametric) panel and (nonparametric) cross-sectional approaches. Event-study plots in the first two panels show no evidence of any effect of DA partisanship on the probability that defendants re-appear on cases within 1 or 2 years of initial arrest. Moreover, relatively tight confidence intervals and point estimates very near to zero indicate a true null effect. Similarly, cross-sectional aggregate data from four years post-election show no discontinuity in recidivism rates among closely-decided jurisdiction-elections, supporting the same conclusion as our parametric panel evidence.

We formally estimate the causal impact of DA partisanship on recidivism using Equations 2.1 and 2.2; our results appear in Table 7. Our panel RD point estimates confirm that Democratic DAs have a null impact on the probability of defendant recidivism. These results in columns 2 through 4 amount to zero effects: the 95 percent confidence intervals on our preferred estimates in column 2 imply that we can rule out an increase in 1-year re-offense probability greater than 0.6 percentage points (2 percent of the sample mean) and an increase in 2-year re-offense probability greater than 0.8 percentage points (also 2 percent of the sample mean). Cross-sectional RD estimates in the remaining columns point to sizeable declines in recidivism following a Democratic DA's election. Imprecise coefficients point to a decline in 1-year recidivism of between 1.1 and 1.5 percentage points (about 4 percent of the sample mean). Interestingly, we find a very precise and economically large reduction of 5.7

our setting, though, looking at outcomes downstream of arrest—such as incarceration—does not make sense, since they will just reflect the impact of DA partisanship, which we have already shown to be substantial.

[40]Our sample size attenuates rapidly when we look at re-offense rates over longer time horizons. Agan et al. (2022), for one, focus on 2-year recidivism, and so we lean into this outcome as a benchmark against the rest of the literature.

percentage points (14 percent) in 2-year recidivism in the fourth year following a Democratic DA's election. However, we caution that the visual evidence of this effect (in Figure 2.5) is wanting, and so we hesitate to accept this estimate at face value.

Still, these results suggest that, at a minimum, Democratic DAs do not encourage greater re-offense rates than their Republican counterparts, and potentially that they reduce defendant recidivism several years after their election. While we reiterate that these estimates do not capture the causal effect of case dismissal, we view these null results as part of a growing body of work that finds diverting or dismissing criminal cases discourages defendants from committing future crimes. Our findings provide suggestive evidence consistent with the conclusion that higher dismissal and lower incarceration rates do not encourage criminal behavior, even if we cannot identify these effects directly.

2.6 External Validity

Heretofore, we have described only the local average treatment effect (LATE) of Democratic DAs on case and defendant outcomes. The estimates we have reported identify the causal effect of DA partisanship among those officials elected in the most competitive elections. Given the prevalence of uncontested DA races, identifying the impact of the average DA's partisanship is of particular policy importance, and there is ample reason to doubt whether our RD approach delivers that average treatment effect (ATE), not least because of the disparities between contested and uncontested jurisdictions we document in Table 3. But delivering a plausibly causal estimate of the average effect of a Democratic DA is challenging: we would need a valid instrument for DA partisanship across all jurisdictions, contested and uncontested, which we do not have.

In the absence of such an instrument, we employ a matching design, which draws on our case-level data to create "matched" samples of otherwise similar cases from similar jurisdictions prosecuted by Democratic and Republican DAs. Our approach combines a typical propensity score design to match jurisdictions with exacting matching to compare similar cases within those jurisdictions. Functionally, we use a logit model to determine the likelihood that a given jurisdiction has a Democratic prosecutor based on its demographic characteristics (those given in Table 3: population, income per capita, nonwhite population share, and caseload per capita). Within matched jurisdictions with similar propensity scores, we then require exact matches of treated and control cases based on key characteristics—the defendant and case covariates listed in Table 2, in addition to the year of case filing, and the lead (most serious) charge on a case.[41] In so doing, we aim to create matched sample of comparable treated and control cases, minimizing the differences in confounding jurisdiction-level and case-level factors in order to isolate the effect of DA partisanship.

2.6.1 Evaluating the Matching Design

We conduct three preliminary analyses to assess the viability of our matching design. The key concern with any matching approach is that the chosen mix of covariates used to construct the matched sample are not sufficient to remove selection bias that distinguishes treated and untreated units. Helpfully, in our setting, we have a benchmark that we can use to assess the performance of our matching approach: our well-identified RD estimates. Intuitively, assuming that our RD estimates capture the "true" causal impact of DA partisanship on case outcomes—locally within the RD subsample—then we would expect a valid matching estimator to yield comparable results when applied to the RD sample. Specifically, we attempt to mimic the multi-period cross-sectional RD approach from column 7 of Table 5 and column 6 of Table 7, which shares key features with our matching design (namely, both make within-year and within-state comparisons).

These matching results appear in the third column of Table 9, alongside the benchmark cross-sectional RD estimates in column 2. Overall, within our RD sample, we recover

[41] In most state datasets, the lead charge is denoted as the first charge on the case; when the data do not indicate the lead charge, we assume it to be the charge with the longest expected incarceration sentence.

matching estimates that are virtually identical to our cross-sectional results, at least when we consider dismissal, incarceration, and recidivism rates. Estimates on sentence length, however, differ markedly: our matching approach recovers an effect less than half as large as our RD specification. As such, we treat our remaining matching estimates of sentencing effects with caution. On the other hand, the similarity across estimates for the remaining outcomes bolsters our faith in the matching estimates outside of the RD sample.

As a second assessment of our matching design, we mimic our panel RD evidence in Figures 2.3 and 2.4 by plotting event-study estimates using our matched sample. Ideally, our matching framework should produce similar event-study coefficients as our panel RD specification, even without the RD controls. Those plots appear in the Appendix, and show that, within the RD sample, our matching and RD estimates track one another reasonably well.

Finally, to confirm that our matching design actually yields similar treatment and control groups, we compare our matched treated and control observations from our full case sample (shown in columns 6 and 7 of Table 9). As expected, we find that jurisdiction and case characteristics are balanced across matched cases prosecuted by DAs of opposing parties. Taken together, our preliminary matching exercises provide some reassurance that our matching design can replicate most of our main RD findings, and that our mix of covariates represents a reasonable basis for creating a matched sample.

2.6.2 Matching Results

In column 5 of Table 9, we show matching estimates of the effect of DA partisanship using our full sample of cases, which we argue approximates the average effect a Democratic DA has on case outcomes. We find that the average Democratic DA is 6.7 percentage points (18 percent) more likely to dismiss a given case, and 8.5 percentage points (18 percent) less likely to impose an incarceration sentence. Taken at face value, the average Democratic DA appears to seek 18 percent shorter incarceration sentences as well, although we caveat that this result may be suspect, given the concerns above about the reliability of our matching approach when it comes to examining sentence length. Altogether, our matching results indicate that the average effect of DA partisanship on case outcomes is smaller than the marginal effect, though still economically meaningful. Furthermore, the matching estimator recovers null effects on re-offense rates, suggesting that neither the average nor the marginal Democratic DA increases the likelihood of defendant recidivism.

2.7 Concluding Discussion

Elected district attorneys hold considerable sway over local court systems, opening the door for political partisanship to shape the implementation of justice. Understanding the degree to which DAs drive prosecution and incarceration rates provides valuable context to ongoing discussions surrounding the punitiveness and efficacy of the American criminal justice system. To our knowledge, this study provides the first causal analysis of how DA partisanship affects prosecution and incarceration rates, as well as defendant recidivism. Unlike prior work, which considers the consequences of non-prosecution, we evaluate a critical determinant of non-prosecution, shedding new light on the political economy of criminal justice in the process.

We compile a dataset of over 10 million criminal cases, which we then use to evaluate the impact of district attorneys on case outcomes. Applying a close-elections difference-in-differences design, we find that the marginal Democratic DA is 8 percentage points (26 percent) more likely to dismiss (decline to prosecute) criminal cases than their Republican counterpart. Likewise, cases handled by Democratic DAs are 9.7 percentage points (17 percent) less likely to result in incarceration, and receive 32 percent shorter incarceration sentences on average. At the same time, we do not find any evidence that electing Democratic DAs raises re-offense or local crime rates. These findings substantiate anecdotal evidence that DA partisanship matters, and in fact causes prosecutorial policies to diverge sharply across jurisdictions. Equally, our study provides unique evidence of how the institutional structure of the justice system—and its reliance on elected prosecutors—can generate more prosecutions and higher incarceration rates without raising measures of criminal behavior.

Our work builds on prior research by highlighting the critical role DA politics plays in determining dismissal rates and, by extension, which defendants experience the harsh economic consequences of accruing a criminal record. As an empirical consensus develops that high prosecution and incarceration have social costs and sometimes elusive benefits, evaluating the roots of more punitive criminal justice policies could highlight avenues for change. This paper contributes to that effort, documenting the critical role DAs can play in supporting harsher, though not necessarily more effectual, court systems. Our results suggest that any future efforts to make the criminal justice system less punitive must overcome the partisan divide among public prosecutors, and run straight through the ballot box.

2.8 Figures

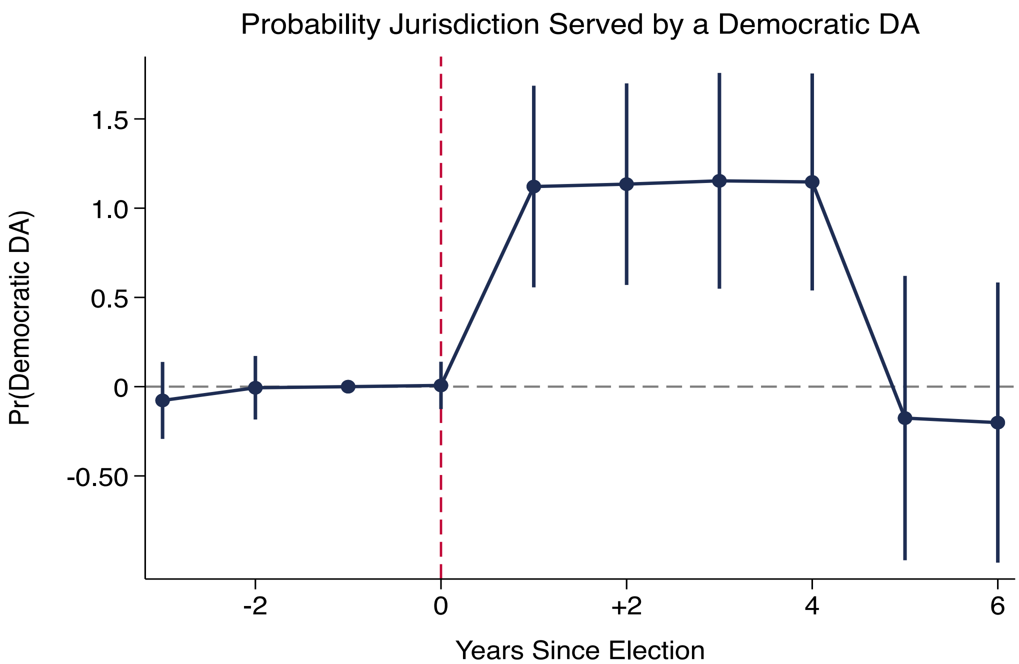


Figure 2.1: The figure plots panel regression estimates showing the impact of a narrow Democratic DA win on the probability that a Democratic DA serves in office in the years around the election. The specification is Equation 2.3. All coefficients are relative to the year prior to the election (period -1). The sample includes 49 competitive district attorney races with at least one Democrat and one Republican running that were decided by 7.1 percentage points or less. Standard errors are clustered at the election level.

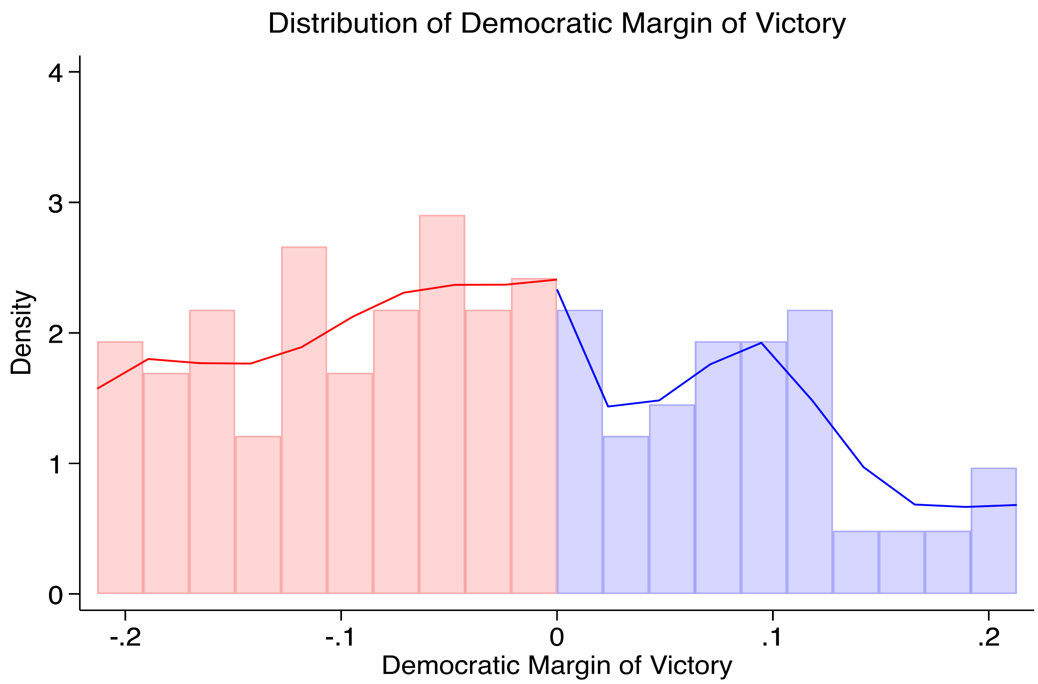


Figure 2.2: The sample includes 134 competitive district attorney races with at least one Democrat and one Republican running that were decided by 20 percentage points or less. "Democratic margin of victory" refers to the difference between the top-performing Democrat's vote share and the top-performing Republican's vote share in the election.

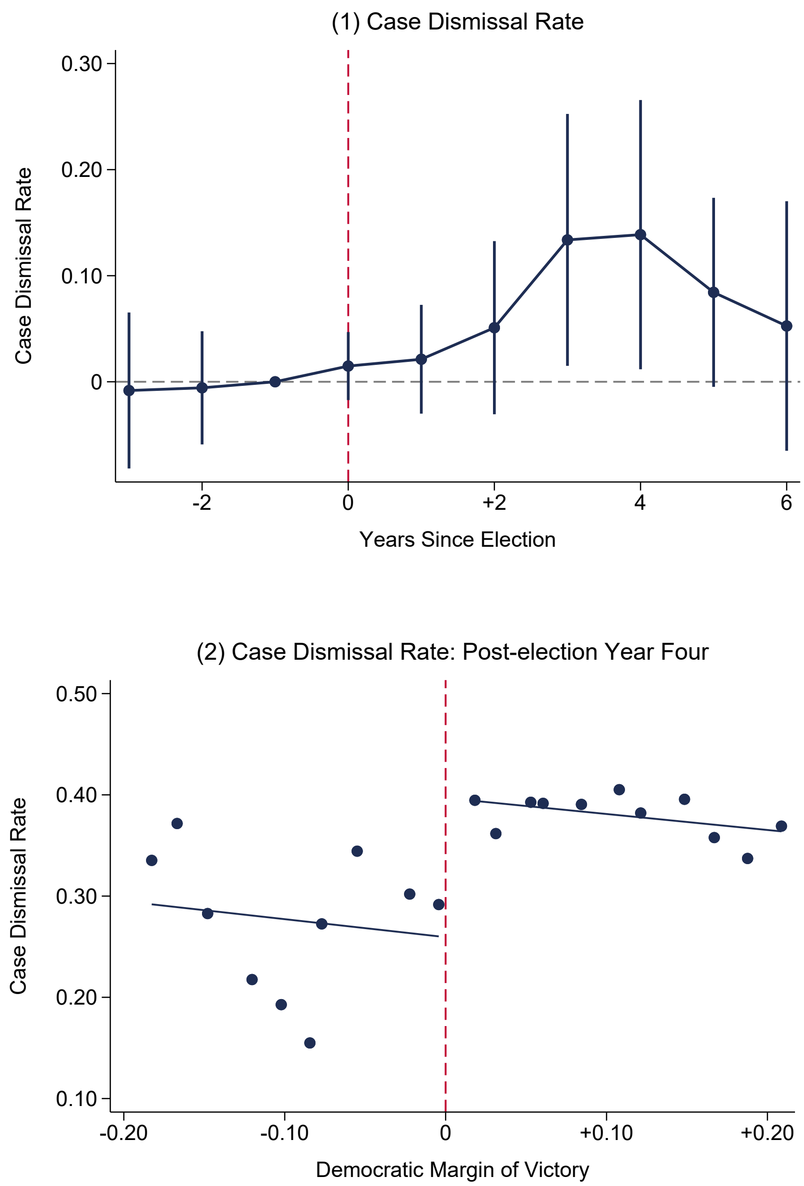


Figure 2.3: Panel 1 plots panel RD estimates showing the impact of a narrow Democratic DA election win on the probability of case dismissal by year relative to the election. The specification is Equation 2.3. All coefficients are relative to the year prior to the election (period -1). The sample includes cases filed in jurisdiction-elections decided by 7.1 percentage points or less (N= 4,160,428 cases). Vertical bars denote 95 percent confidence intervals using robust standard errors clustered at the jurisdiction-election level. Panel 2 presents a bin scatter plot of case jurisdiction-level dismissal rates in the fourth year post-election by the Democratic candidate's margin of victory (or loss). The sample contains 134 jurisdiction-elections.

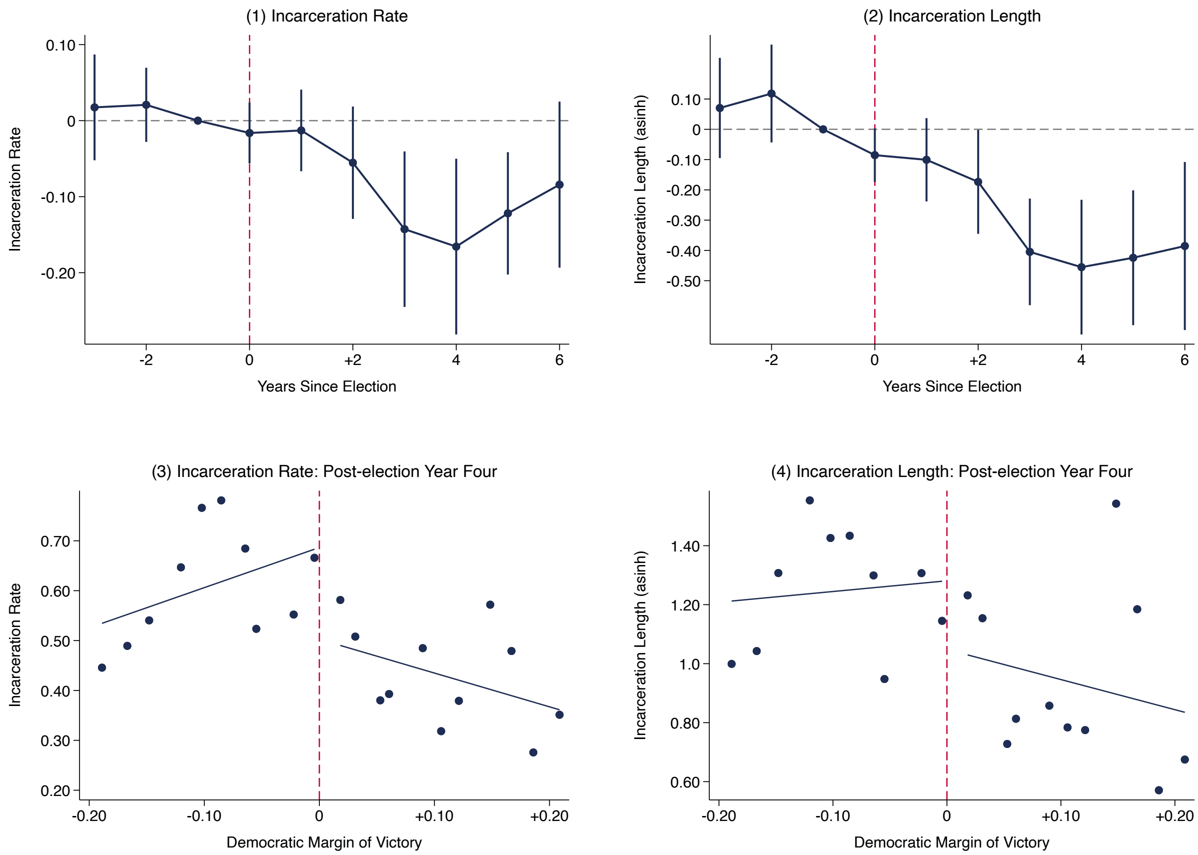


Figure 2.4: Panels 1 and 2 plot panel RD estimates showing the impact of a narrow Democratic election win on the probability of incarceration (panel 1) and incarceration length (panel 2) by year relative to the election. The specification is Equation 2.3. All coefficients are relative to the year prior to the election (period -1). The sample includes cases filed in jurisdiction-elections decided by 7.1 percentage points or less (N= 4,160,428). Vertical bars denote 95 percent confidence intervals using robust standard errors clustered at the jurisdiction-election level. Panels 3 and 4 shows bin scatter plots of jurisdiction-wide incarceration rates and mean sentence lengths in the fourth year post-election by the Democratic candidate's margin of victory (or loss). The sample contains 134 jurisdiction-elections.

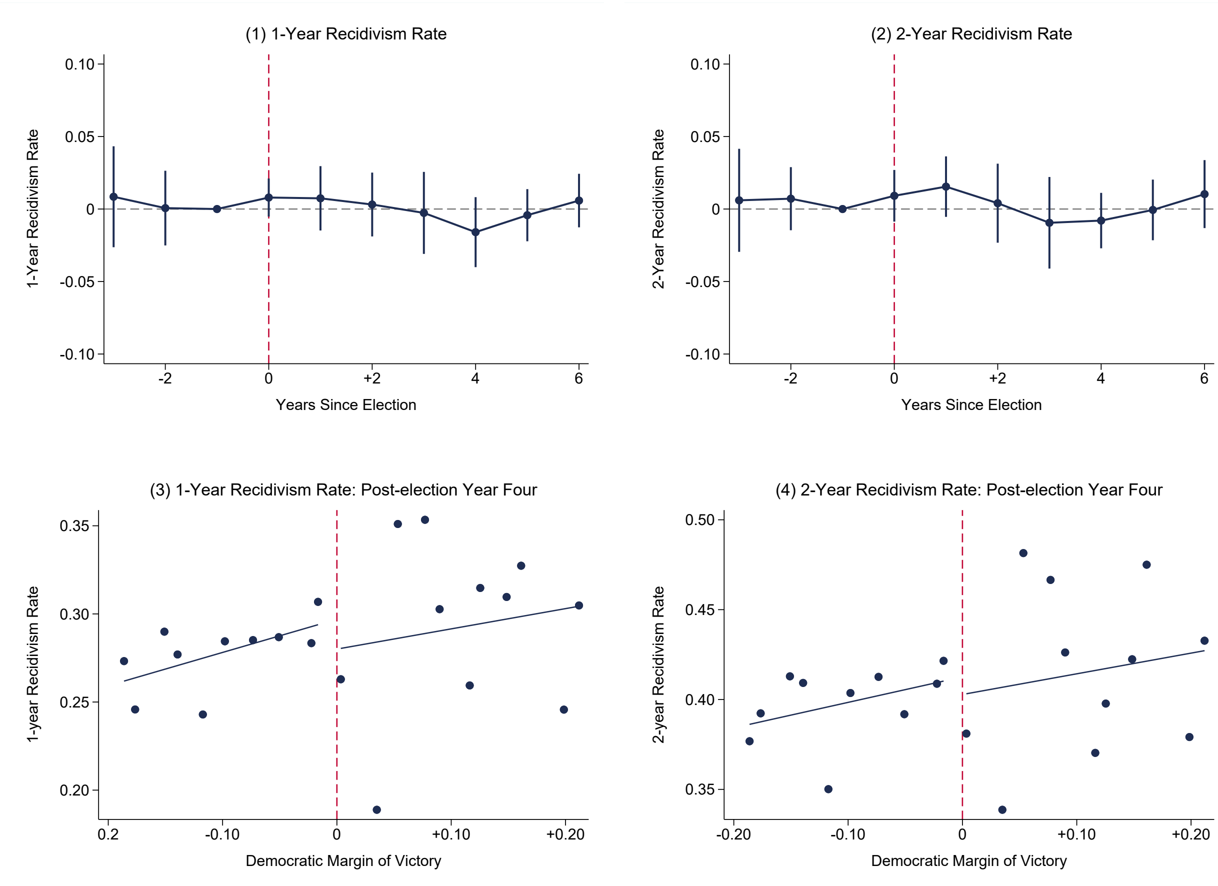


Figure 2.5: Panels 1 and 2 plot panel RD estimates showing the impact of a narrow Democratic election win on the probability of defendant recidivism within 1 year of initial arrest (panel 1) and the probability of defendant recidivism within 2 years of initial arrest (panel 2), by year relative to the election. The specification is Equation 2.3. All coefficients are relative to the year prior to the election (period -1). The sample includes cases filed in jurisdiction-elections decided by 7.1 percentage points or less (N=3,474,473 for panel 1, and 3,322,341 for panel 2). Vertical bars denote 95 percent confidence intervals using robust standard errors clustered at the jurisdiction-election level. Panels 3 and 4 shows bin scatterplots of jurisdiction-wide incarceration rates and mean sentence lengths in the fourth year post-election by the Democratic candidate's margin of victory (or loss). The sample contains 87 jurisdiction-elections.

2.9 Tables

Table 1: Summary of District Attorney Elections

	N	Mean	Median	Std Dev	Min	Max
I. Election Characteristics						
Election year	1,311	2008	2008	6	1996	2017
# of Candidates	1,311	1.21	1	0.43	1	4
# of Democrats	1,311	0.54	1	0.50	0	2
Contested Election?	1,311	0.19	0	0.40	0	1
II. Election Outcomes						
Democrat Won?	1,311	0.43	0	0.50	0	1
Republican Won?	1,311	0.55	1	0.50	0	1
Election Margin	255	0.18	0.14	0.17	0	1
Dem-Rep Margin	197	-0.06	-0.05	0.19	-0.55	0.36

The table summarizes outcomes from 1,311 district attorney elections held in Arkansas, Colorado, Kentucky, North Carolina, Texas, and Virginia between 1996 and 2017. The election margin describes the difference between the first- and second-place candidates' vote shares, irrespective of their political parties, whereas the "Dem-Rep Margin" describes the difference in vote share between the leading Democratic and Republican candidates. Sample sizes vary because not all DA elections are contested, and not all contested elections have both a Democratic and a Republican candidate.

Table 2: Criminal Case-level Descriptive Statistics by State

		State of Case Filing				
	All Cases	Arkansas	Colorado	North Carolina	Texas	Virginia
	(1)	(2)	(3)	(4)	(5)	(6)
I. Case Outcomes						
Case Dismissed?	0.37	0.19	0.29	0.58	0.31	0.43
	(0.48)	(0.39)	(0.45)	(0.49)	(0.46)	(0.49)
Incarceration?	0.47	0.56	0.37	0.15	0.62	0.28
	(0.50)	(0.50)	(0.48)	(0.36)	(0.48)	(0.45)
Incarceration Length (asinh)	0.99	2.08	0.91	0.21	1.31	0.55
	(1.50)	(2.26)	(1.51)	(0.71)	(1.63)	(1.03)
N:	*9,787,182*	*79,729*	*1,376,024*	*1,616,202*	*5,519,997*	*1,195,230*
1-year Recidivism	0.30	—	—	0.38	0.28	—
	(0.46)	—	—	(0.49)	(0.45)	—
N:	*6,896,017*	—	—	*1,474,430*	*5,421,587*	—-
2-year Recidivism	0.41	.	—	0.48	0.40	—
	(0.49)	—	—	(0.50)	(0.49)	—
N:	*6,352,006*	—	—	*1,205,969*	*5,146,037*	—-
II. Defendant Characteristics						
Age	32.33	35.60	32.47	33.29	31.97	—
	(11.09)	(14.51)	(11.21)	(11.81)	(10.76)	—
N:	*8,591,952*	*79,729*	*1,376,024*	*1,616,202*	*5,519,997*	—
Female	0.25	0.26	0.23	0.29	0.23	0.30
	(0.43)	(0.44)	(0.42)	(0.45)	(0.42)	(0.46)
Nonwhite	0.54	0.30	0.20	0.50	0.66	0.41
	(0.50)	(0.46)	(0.40)	(0.50)	(0.47)	(0.49)
III. Case Characteristics						
# of Charges	1.70	2.92	2.32	2.19	1.43	1.48
	(1.91)	(3.64)	(1.58)	(3.61)	(0.99)	(1.69)
Felony Offense?	0.34	0.27	0.42	0.22	0.36	0.32
	(0.47)	(0.44)	(0.49)	(0.41)	(0.48)	(0.47)
Property Offense?	0.35	0.37	0.31	0.45	0.34	0.34
	(0.48)	(0.48)	(0.46)	(0.50)	(0.47)	(0.47)
Violent Offense?	0.26	0.16	0.30	0.35	0.24	0.16
	(0.44)	(0.36)	(0.46)	(0.48)	(0.43)	(0.36)
Drug Offense?	0.27	0.24	0.21	0.24	0.31	0.22
	(0.44)	(0.42)	(0.41)	(0.43)	(0.46)	(0.41)
Traffic Offense?	0.07	0.19	0.25	0.03	0.04	0.03
	(0.25)	(0.40)	(0.43)	(0.17)	(0.20)	(0.17)
Other Offense?	0.38	0.52	0.64	0.30	0.34	0.43
	(0.49)	(0.50)	(0.48)	(0.46)	(0.47)	(0.49)
N:	*9,787,182*	*79,729*	*1,376,024*	*1,616,202*	*5,519,997*	*1,195,230*

[1] Each cell reports the mean of the variable in the left-hand column, with standard deviations in parentheses. Column 1 describes all cases in our dataset (N=9,787,182). The remaining columns describe cases by the state in which they were filed. Empty cells reflect missing recidivism data and information on defendant age. Sample sizes vary within columns due to missing recidivism data. See the text for more detail on data missingness and sample construction.

Table 3: Comparing Democratic- and Republican-led DA Jurisdictions

		Democratic DAs		Republican DAs	
	Full Sample	Uncontested Election	Contested Election	Contested Election	Uncontested Election
	(1)	(2)	(3)	(4)	(5)
I. Jurisdiction Characteristics					
Population	171,319	117,029	342,808	376,969	133,534
	(391,646)	(186,787)	(637,433)	(804,165)	(225,494)
Share Nonwhite	0.24	0.25	0.24	0.24	0.24
	(0.19)	(0.22)	(0.22)	(0.18)	(0.17)
N:	*3,874*	*1,335*	*252*	*498*	*1,735*
Income per Capita ($2016)	38,888	34,877	40,597	40,104	41,369
	(10,516)	(10,031)	(11,159)	(10,314)	(10,010)
N:	*3,574*	*1,229*	*234*	*459*	*1,610*
II. Court Outcomes					
Annual Caseload	2,526	1,922	5,267	4,987	1,947
	(5,311)	(3,026)	(8,902)	(10,648)	(2,902)
Caseload per 1,000 Pop	17.2	17.2	18.6	17.9	16.7
	(10.1)	(10.5)	(9.7)	(11.2)	(9.3)
Case Dismissal Rate	0.37	0.41	0.39	0.35	0.35
	(0.12)	(0.13)	(0.11)	(0.13)	(0.12)
Incarceration Rate	0.44	0.40	0.39	0.45	0.48
	(0.19)	(0.18)	(0.17)	(0.20)	(0.20)
Avg. Sentence Length (asinh)	1.08	1.00	0.94	1.13	1.18
	(0.62)	(0.57)	(0.60)	(0.72)	(0.61)
N:	*3,874*	*1,335*	*252*	*498*	*1,735*

The data describe the characteristics and judicial outcomes among DA jurisdiction-years between 2000 and 2019. Column 1 reports the mean of the given variable across all jurisdiction-years (N=3,874). Columns 2 and 3 describe jurisdiction-years with serving Democratic DAs elected in uncontested (column 2) and contested (column 3) elections, while columns 4 and 5 describe jurisdiction-years with serving Republican DAs elected in uncontested (column 4) and contested (column 5) elections. Standard deviations appear in parentheses. Population data come from the Census intercensal estimates, while income per capita data come from the Bureau of Economic Analysis (BEA). Sample sizes vary within column because the BEA income data do not include many localities in Virginia. Sample sizes vary across columns because some DA election winners represent third parties, while others do not have an identified partisan affiliation.

Table 4: RD Validity: Election-year Differences in Jurisdiction and Case Characteristics

	Control Mean	Balance Estimate
	(1)	(2)
I. Jurisdiction Characteristics		
Population	807,711	57,216
	(1,310,659)	(653,494)
Caseload per 1,000 Pop	15.5	-1.1
	(6.3)	(3.2)
N:	*29*	*49*
Share Nonwhite	0.255	-0.006
	(0.212)	(0.124)
N:	*29*	*47*
Income per Capita ($2016)	44,299	5,278
	(11,351)	(6,253)
N:	*27*	*43*
II. Case Characteristics		
Defendant Age	32.2	0.2
	(10.9)	(0.4)
N:	*307,769*	*428,855*
Defendant Female?	0.232	-0.016
	(0.422)	(0.014)
Defendant Nonwhite?	0.680	0.019
	(0.466)	(0.073)
# of Charges	1.49	-0.01
	(1.11)	(0.15)
Felony Offense?	0.373	0.021
	(0.483)	(0.027)
Property Offense?	0.338	0.027
	(0.473)	(0.031)
Violent Offense?	0.226	0.040
	(0.418)	(0.030)
Drug Offense?	0.308	-0.051
	(0.462)	(0.046)
Traffic Offense?	0.062	-0.005
	(0.241)	(0.023)
Other Offense?	0.352	0.022
	(0.478)	(0.038)
N:	*313,870*	*444,159*

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The sample includes election-year jurisdictions (N=49) and cases (N=444,160) for which the upcoming election is decided by 7.1 percentage points or less. Column 1 reports the mean of the outcome variable in the left-hand column among jurisdictions in which Republican candidates win the election. Standard deviations appear in parentheses. Column 2 presents cross-sectional estimates of the Democratic DA effect, following Equation 2.1. Robust standard errors clustered at the jurisdiction-election level appear in parentheses. Sample sizes vary within columns due to missing jurisdiction demographic and defendant age data, as discussed in the text.

Table 5: How Do Democratic DAs Affect Case Dispositions? Panel and Cross-sectional Estimates

	Sample Mean	Panel Estimates			Cross-sectional Estimates		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
I. Probability of Dismissal							
Case Dismissal	0.319	0.083**	0.075**	0.074**	0.147	0.080*	0.101**
	(0.466)	(0.035)	(0.033)	(0.034)	(0.095)	(0.042)	(0.039)
II. Incarceration Outcomes							
Incarceration	0.579	-0.097***	-0.092***	-0.098***	-0.084	-0.079**	-0.113***
	(0.493)	(0.032)	(0.030)	(0.03)	(0.104)	(0.037)	(0.041)
Incarceration Length (asinh)	1.134	-0.318***	-0.364***	-0.385***	0.010	-0.224**	-0.231***
	(1.521)	(0.062)	(0.057)	(0.063)	(0.189)	(0.093)	(0.105)
N:	*8,303,796*	*3,490,129*	*3,490,129*	*4,160,428*	*49*	*49*	*282*
Year FEs		Y	Y	Y	N	N	Y
Jurisdiction FEs		Y	Y	Y	N	N	Y
Defendant/Case Covariates		N	Y	Y	—	—	—
Period(s) Included		[-3,4]	[-3,4]	[-3,6]	4	[-3,4]	[1,6]
Unit of Observation		Case	Case	Case	Election	Election	Election-period

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The sample includes criminal cases filed between pre-election period -3 and post-election period +6 in jurisdiction-elections decided by 7.1 percentage points or less. Column 1 provides the sample mean of each outcome in the left-hand column. Standard deviations appear in parentheses. Columns 2-4 report panel estimates of the post-election effect of a Democratic DA victory over the four years following an election (columns 2 and 3) and over the six years following an election (column 4). The panel specification is Equation 2.2. Columns 5-7 report cross-sectional estimates of the impact of a Democratic DA victory on jurisdiction-wide mean outcomes the fourth year following an election (column 5), in the difference between pre- and post-election outcomes (column 6), and over the six years post-election (column 7). The cross-sectional RD specification is Equation 2.1. All cross-sectional specifications employ jurisdiction population weights. All panel and cross-sectional specifications include linear controls for the Democratic candidate's margin of victory (or loss) and its interaction with an indicator for whether a Democratic candidate won the election. Additional defendant and case covariates in the panel specification include indicators for whether a defendant is nonwhite or female, whether the case includes felony, property, violent, traffic, drug, or other charges, and the total number of charges on the case. Robust standard errors clustered at the jurisdiction-election level appear in parentheses.

Table 6: Robustness of Panel Estimates to Alternative Specifications and Samples

	Baseline Specification (1)	20 Percent Narrower BW (2)	100 Percent Wider BW (3)	Quadratic Polynomial (4)	"Donut" of ± 1pp (5)
Case Dismissal	0.083**	0.066***	0.050**	0.089**	0.150***
	(0.035)	(0.022)	(0.020)	(0.039)	(0.048)
Incarceration	-0.097***	-0.077***	-0.042*	-0.105***	-0.141***
	(0.032)	(0.020)	(0.025)	(0.037)	(0.052)
Incarceration Length (asinh)	-0.318***	-0.282***	-0.218***	-0.340***	-0.392***
	(0.062)	(0.058)	(0.056)	(0.075)	(0.101)
N:	*3,490,121*	*2,721,254*	*5,527,517*	*3,490,121*	*2,688,401*
Year FEs	Y	Y	Y	Y	Y
Jurisdiction FEs	Y	Y	Y	Y	Y
Defendant/Case Covariates	N	N	N	N	N
Periods Included	[-3,4]	[-3,4]	[-3,4]	[-3,4]	[-3,4]
"Close" Election Margin	± 7.1pp	± 5.7pp	± 14.2pp	± 7.1pp	± 7.1pp

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The sample includes criminal cases filed between pre-election period -3 and post-election period +4. Each cell reports a point estimate capturing the post-election effect of a Demcoratic DA victory. Our baseline specification is Equation 2.2, estimated among cases filed in jurisdiction-elections decided by less than our optimal bandwidth of 7.1 percentage points. Columns 2, 3, and 5 present estimates from Equation 2.2 estimated on alternative samples described in the header. Column 4 presents results from a modified version of Equation 2.2 that includes a quadratic control for the election margin, interacted with an indicator for whether a Democrat won the election. Robust standard errors clustered at the jurisdiction-election level appear in parentheses.

Table 7: How Do Democratic DAs Affect Recidivism Rates? Panel and Cross-sectional Estimates

	Sample Mean	Panel Estimates			Cross-sectional Estimates		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1-year Recidivism	0.290	-0.004	0.002	-0.001	-0.011	-0.011	-0.015*
	(0.454)	(0.005)	(0.005)	(0.004)	(0.041)	(0.028)	(0.008)
N:	*3,474,473*	*2,920,924*	*2,920,924*	*3,474,473*	*31*	*28*	*173*
2-year Recidivism	0.411	-0.002	0.002	-0.001	-0.057***	-0.010	-0.022*
	(0.492)	(0.005)	(0.005)	(0.005)	(0.015)	(0.034)	(0.011)
N:	*3,322,341*	*2,815,563*	*2,815,563*	*3,322,341*	*28*	*28*	*163*
Year FEs	Y	Y	Y	N	N	Y	
Jurisdiction FEs	Y	Y	Y	N	N	Y	
Defendant/Case Covariates	N	Y	Y	—	—	—	
Period(s) Included	[-3,4]	[-3,4]	[-3,6]	4	[-3,4]	[1,6]	
Unit of Observation	Case	Case	Case	Election	Election	Election-period	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The sample includes criminal cases filed between pre-election period -3 and post-election period +6 in jurisdiction-elections decided by 7.1 percentage points or less. Column 1 provides the sample mean of each outcome in the left-hand column. Standard deviations appear in parentheses. Columns 2 and 3 report panel estimates of the post-election effect of a Democratic DA victory over the four years following an election. The panel specification is Equation 2.2. Columns 5-7 report cross-sectional estimates of the impact of a Democratic DA victory on jurisdiction-wide mean outcomes in the fourth year following an election (column 5), in the difference between pre- and post-election outcomes (column 6), and over the six years post-election (column 7). The cross-sectional specification is Equation 2.1. All cross-sectional specifications employ jurisdiction population weights. All panel and cross-sectional specifications include a linear control for the Democratic candidate's margin of victory (or loss) and its interaction with an indicator for whether a Democratic candidate won the election. Additional defendant and case covariates are described below Table 5. Robust standard errors clustered at the jurisdiction-election level appear in parentheses.

Table 8: Heterogeneity in Democratic DA Treatment Effects: Panel Estimates

	All	Defendant Characteristics					
		Race/Ethnicity		Gender		Age	
	Cases	Nonwhite	White	Female	Male	Age≤30	Age>30
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Case Dismissal	0.083**	0.093***	0.078**	0.095**	0.077**	0.082**	0.080**
	(0.031)	(0.030)	(0.038)	(0.045)	(0.032)	(0.038)	(0.032)
	0.315	*0.306*	*0.334*	*0.398*	*0.290*	*0.326*	*0.305*
Incarceration	-0.097***	-0.118***	-0.071**	-0.113**	-0.090***	-0.109***	-0.086***
	(0.032)	(0.027)	(0.034)	(0.043)	(0.029)	(0.036)	(0.029)
	0.602	*0.650*	*0.503*	*0.511*	*0.629*	*0.597*	*0.607*
Incarceration Length (asinh)	-0.318***	-0.348***	-0.280***	-0.247***	-0.330***	-0.340***	-0.295***
	(0.062)	(0.054)	(0.077)	(0.071)	(0.059)	(0.063)	(0.067)
	1.149	*1.211*	*1.023*	*0.808*	*1.253*	*1.085*	*1.211*
N:	*3,490,121*	*2,347,973*	*1,142,148*	*814,374*	*2,675,747*	*1,711,906*	*1,778,215*

	Case Characteristics						
	Case Severity		Violence		Types of Offense		
	Felony	Misdemeanor	Violent	Nonviolent	Property	Drug	Other
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Case Dismissal	0.105*	0.072**	0.058*	0.092**	0.096***	0.067*	0.101**
	(0.054)	(0.032)	(0.034)	(0.037)	(0.034)	(0.040)	(0.041)
	0.283	*0.334*	*0.390*	*0.292*	*0.285*	*0.273*	*0.316*
Incarceration	-0.108*	-0.091**	-0.067**	-0.107***	-0.110***	-0.061*	-0.134***
	(0.056)	(0.038)	(0.031)	(0.033)	(0.031)	(0.035)	(0.044)
	0.650	*0.574*	*0.536*	*0.623*	*0.646*	*0.658*	*0.554*
Incarceration Length (asinh)	-0.648***	-0.254***	-0.329***	-0.315***	-0.391***	-0.246***	-0.0376***
	(0.150)	(0.047)	(0.074)	(0.063)	(0.071)	(0.077)	(0.068)
	2.116	*0.594*	*1.419*	*1.064*	*1.178*	*1.301*	*0.990*
N:	*1,273,577*	*2,216,544*	*839,378*	*2,650,743*	*1,189,213*	*1,004,548*	*1,264,743*
Year FEs	Y	Y	Y	Y	Y	Y	Y
Jurisdiction FEs	Y	Y	Y	Y	Y	Y	Y
Defendant/Case Covariates	N	N	N	N	N	N	N
Periods Included	[-3,4]	[-3,4]	[-3,4]	[-3,4]	[-3,4]	[-3,4]	[-3,4]

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The sample includes criminal cases filed between pre-election period -3 and post-election period +4 in jurisdiction-elections decided by 7.1 percentage points or less. Each cell reports panel estimates of the effect of a Democratic DA victory, following Equation 2.2. Each coefficient comes from a separate regression, estimated among the subsample of cases described in the column header. Robust standard errors clustered at the jurisdiction-election level appear in parentheses. The subsample mean of the outcome variable in the left-hand column appears in italics below the standard error.

Table 9: Assessing the Average Effect of DA Partisanship: Matching Estimates

	Close-elections Sample			All Cases	
	Sample Mean	Regression Estimates	Matching Estimates	Sample Mean	Matching Estimates
	(1)	(2)	(3)	(4)	(5)
Case Dismissal	0.363	0.101**	0.100**	0.367	0.067***
	(0.481)	(0.039)	(0.042)	(0.482)	(0.022)
Incarceration	0.556	-0.113***	-0.112**	0.466	-0.085**
	(0.497)	(0.041)	(0.054)	(0.499)	(0.034)
Incarceration Length (asinh)	1.057	-0.231***	-0.099	0.988	-0.179***
	(1.487)	(0.081)	(0.098)	(1.504)	(0.060)
N:	*2,453,178*	*282*	*1,858,927*	*9,787,182*	*7,979,231*
1-year Recidivism	0.291	-0.007	-0.006	0.302	0.006
	(0.454)	(0.007)	(0.009)	(0.459)	(0.007)
N:	*2,044,758*	*173*	*1,706,426*	*6,896,017*	*6,152,973*
2-year Recidivism	0.410	-0.014	-0.000	0.414	0.006
	(0.492)	(0.009)	(0.008)	(0.492)	(0.008)
N:	*1,892,626*	*163*	*1,563,230*	*6,352,006*	*5,686,948*

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1-3 include observations filed between post-election periods +1 and +6 in jurisdiction-elections decided by 7.1 percentage points or less. Column 1 provides the sample mean of the each outcome in the left-hand column. Standard deviations appear in parentheses. Column 2 reproduces cross-sectional RD estimates from column 7 of Table 5 and column 6 of Table 7. The specification in Equation 2.1, estimated on jurisdiction-election-period data with jurisdiction and year fixed effects. Column 3 presents matching results using the same sample of cases as in column 1, following the approach described in the text. Columns 4 and 5 include all cases in our sample. Column 4 provides the sample mean, while column 5 presents matching estimates following the same design as in column 3. The differences in sample sizes between columns 1 and 3, as well as between columns 4 and 5, represent observations that were not matched in our matching process. Robust standard errors clustered at the jurisdiction-election level appear in parentheses in columns 2, 3, and 5.

2.10 Appendix

Table 10: Criminal Case-level Descriptive Statistics by DA Partisanship

		DA Partisanship	
	Full Sample	Democrat	Republican
	(1)	(2)	(3)
I. Case Outcomes			
Case Dismissal	0.37	0.43	0.33
	(0.48)	(0.49)	(0.47)
Incarceration	0.47	0.39	0.52
	(0.50)	(0.49)	(0.50)
Incarceration Length (asinh)	0.99	0.82	1.11
	(1.50)	(1.40)	(1.56)
N:	9,787,182	3,892,610	5,856,861
1-year Recidivism	0.30	0.31	0.30
	(0.46)	(0.46)	(0.46)
N:	6,896,017	2,661,861	4,234,156
2-year Recidivism	0.41	0.42	0.41
	(0.49)	(0.49)	(0.49)
N:	6,352,006	2,425,282	3,926,724
II. Defendant Characteristics			
Age	32.33	32.41	32.28
	(11.09)	(11.21)	(11.02)
N:	8,591,952	3,286,171	5,305,781
Female?	0.25	0.25	0.25
	(0.43)	(0.43)	(0.43)
Nonwhite?	0.54	0.58	0.51
	(0.50)	(0.49)	(0.50)
III. Case Characteristics			
# of Charges	1.70	1.73	1.68
	(1.91)	(2.22)	(1.68)
Felony?	0.34	0.34	0.34
	(0.47)	(0.47)	(0.47)
Property Offense?	0.35	0.36	0.35
	(0.48)	(0.48)	(0.48)
Violent Offense?	0.26	0.27	0.25
	(0.44)	(0.44)	(0.43)
Drug Offense?	0.27	0.26	0.28
	(0.44)	(0.44)	(0.45)
Traffic Offense?	0.07	0.07	0.07
	(0.25)	(0.25)	(0.26)
Other Offense?	0.38	0.39	0.38
	(0.49)	(0.49)	(0.49)
N:	9,787,182	3,892,610	5,856,861

Each cell reports the mean of the variable in the left-hand column, with standard deviations in parentheses. Column 1 describes all cases in our dataset (N=9,787,182). Column 2 focuses on cases filed in jurisdiction-years in which a Democratic DA held office (N=3,892,610). Column 3 focuses on cases filed in jurisdiction-years in which a Republican DA held office (N=5,856,861). Sample sizes vary within columns because of missing recidivism and defendant age information in particular states. See the text for more detail on data missingness and sample construction.

Table 11: Criminal Case-level Descriptive Statistics Across Samples

	Full Sample	Contested Races	RD Sample
	(1)	(2)	(3)
I. Case Outcomes			
Case Dismissed?	0.37	0.34	0.36
	(0.48)	(0.47)	(0.48)
Incarceration?	0.47	0.56	0.56
	(0.50)	(0.50)	(0.50)
Incarceration Length (asinh)	0.99	1.11	1.06
	(1.50)	(1.52)	(1.49)
N:	*9,787,182*	*5,121,134*	*2,453,178*
1-year Recidivism	0.30	0.29	0.29
	(0.46)	(0.46)	(0.45)
N:	*6,896,017*	*4,238,301*	*2,044,758*
2-year Recidivism	0.41	0.41	0.41
	(0.49)	(0.49)	(0.49)
N:	*6,352,006*	*3,963,392*	*1,892,626*
II. Defendant Characteristics			
Age	32.33	32.40	32.65
	(11.09)	(11.03)	(11.18)
N:	*8,591,952*	*4,887,943*	*2,399,296*
Female	0.25	0.24	0.23
	(0.43)	(0.43)	(0.42)
Nonwhite	0.54	0.64	0.67
	(0.50)	(0.48)	(0.47)
III. Case Characteristics			
# of Charges	1.70	1.54	1.51
	(1.91)	(1.47)	(1.31)
Felony Offense?	0.34	0.35	0.36
	(0.47)	(0.48)	(0.48)
Property Offense?	0.35	0.34	0.35
	(0.48)	(0.48)	(0.48)
Violent Offense?	0.26	0.25	0.26
	(0.44)	(0.43)	(0.44)
Drug Offense?	0.27	0.28	0.28
	(0.44)	(0.45)	(0.45)
Traffic Offense?	0.07	0.05	0.06
	(0.25)	(0.22)	(0.23)
Other Offense?	0.38	0.36	0.37
	(0.49)	(0.48)	(0.48)
N:	*9,787,182*	*5,121,134*	*2,453,178*

Each cell reports the mean of the variable in the left-hand column, with standard deviations in parentheses. Column 1 describes all cases in our dataset. The remaining columns describe cases filed in the six years post-election, by the degree of election competitiveness. The RD sample includes cases filed after elections decided by 7.1 percentage points or less. Sample sizes vary within columns due to missing data on recidivism and defendant age. See the text for more detail on data missingness and sample construction.

Table 12: Heterogeneity in Democratic DA Treatment Effect on Recidivism: Panel RD Estimates

		Defendant Characteristics					
	All	Race/Ethnicity		Gender		Age	
	Cases	Nonwhite	White	Female	Male	Age≤30	Age>30
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1-year Recidivism	-0.004	-0.007	0.005	0.006	-0.006	-0.001	-0.009
	(0.005)	(0.004)	(0.009)	(0.007)	(0.005)	(0.006)	(0.006)
	0.288	*0.279*	*0.318*	*0.249*	*0.300*	*0.307*	*0.269*
N:	*2,920,924*	*2,208,175*	*712,749*	*673,428*	*2,247,496*	*1,485,648*	*1,435,276*
2-year Recidivism	-0.002	-0.011**	0.020*	0.000	-0.002	0.002	-0.010**
	(0.005)	(0.004)	(0.011)	(0.009)	(0.005)	(0.007)	(0.005)
	0.411	*0.400*	*0.444*	*0.357*	*0.426*	*0.439*	*0.381*
N:	*2,815,563*	*2,815,563*	*685,730*	*649,089*	*2,166,474*	*1,438,237*	*1,377,326*

	Case Characteristics						
	Case Severity		Violence		Types of Offense		
	Felony	Misdemeanor	Violent	Nonviolent	Property	Drug	Other
	(9)	(10)	(11)	(12)	(13)	(14)	(15)
1-year Recidivism	0.013	-0.009*	-0.005	-0.003	-0.001	0.007	-0.018***
	(0.012)	(0.005)	(0.007)	(0.006)	(0.010)	(0.010)	(0.005)
N:	*0.252*	*0.308*	*0.212*	*0.312*	*0.346*	*0.290*	*0.261*
	1,058,852	*1,862,072*	*695,777*	*2,225,147*	*1,008,306*	*878,917*	*932,016*
2-year Recidivism	0.017	-0.012*	0.000	-0.004	-0.003	0.013	-0.021**
	(0.012)	(0.006)	(0.006)	(0.011)	(0.006)	(0.010)	(0.009)
	0.389	*0.423*	*0.318*	*0.439*	*0.469*	*0.427*	*0.378*
N:	*1,019,120*	*1,796,443*	*663,877*	*2,151,686*	*974,833*	*849,920*	*895,719*
Year FEs	Y	Y	Y	Y	Y	Y	Y
Jurisdiction FEs	Y	Y	Y	Y	Y	Y	Y
Defendant/Case Covariates	N	N	N	N	N	N	N
Periods Included	[-3,4]	[-3,4]	[-3,4]	[-3,4]	[-3,4]	[-3,4]	[-3,4]

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The sample includes criminal cases filed between pre-election period -3 and post-election period +4 in jurisdiction-elections decided by 7.1 percentage points or less. Each cell reports panel RD estimates of the effect of a Democratic DA victory, following Equation 2.2. Each coefficient comes from a separate regression, estimated among the subsample of cases described in the column header. Robust standard errors clustered at the jurisdiction-election level appear in parentheses. The subsample mean of the outcome variable in the left-hand column appears in italics below the standard error of each estimate. Sample sizes vary within columns due to missing recidivism data.

Table 13: Specification Robustness: Case-level Estimates of Democratic DA Effect

	Case-level Cross-sectional RD				Case-level Panel RD			
	Period 4 only	+ Periods 1 thru 3	+ Jurisdiction FEs	+ Year FEs	+ Period 0	Preferred Specification	+ Covariates	+ Periods 5 and 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Case Dismissed	0.16*	0.10	0.07	0.09**	0.08*	0.08**	0.07**	0.07**
	(0.08)	(0.09)	(0.11)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)
Incarceration	-0.08	-0.00	-0.07	-0.09*	-0.08**	-0.10***	-0.09***	-0.10***
	(0.11)	(0.10)	(0.10)	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)
Incarceration Length (asinh)	-0.00	0.18	-0.25	-0.18	-0.19**	-0.32***	-0.36***	-0.38***
	(0.20)	(0.19)	(0.35)	(0.12)	(0.08)	(0.06)	(0.06)	(0.06)
N:	433,090	1,782,871	1,782,871	1,782,871	2,227,033	3,490,121	3,490,121	4,160,428
1-year Recidivism	-0.04	-0.03**	0.01	-0.01**	-0.01	-0.00	0.00	-0.00
	(0.03)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
N:	342,142	1,491,209	1,491,209	1,491,209	1,863,078	2,920,924	2,920,924	3,474,473
2-year Recidivism	-0.06***	-0.03**	0.00	-0.02***	-0.00	-0.00	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)
N:	295,725	1,385,848	1,385,848	1,385,848	1,757,717	2,815,563	2,815,563	3,322,341
Year FEs	N	N	N	Y	Y	Y	Y	Y
Jurisdiction FEs	N	N	Y	Y	Y	Y	Y	Y
Defendant/Case Covariates	N	N	N	N	N	N	Y	Y
Period(s) Included	4	[1-4]	[1-4]	[1-4]	[0-4]	[-3-4]	[-3-4]	[-3,6]
Unit of Observation	Case	Case	Case	Case	Case	Case	Case	Case

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Each cell comes from a separate regression. The sample consists of cases filed in jurisdiction-elections decided by 7.1 percentage points or less. The specification in columns 1-4 is Equation 2.1, applied to case-level data. The specification in columns 5-8 is Equation 2.2, applied to different samples of election-periods. Sample sizes differ within columns due to missing recidivism data, as discussed in the text. Case-level covariates are described in the footnote to Table 5. Standard errors in parentheses are clustered at jurisdiction-election level.

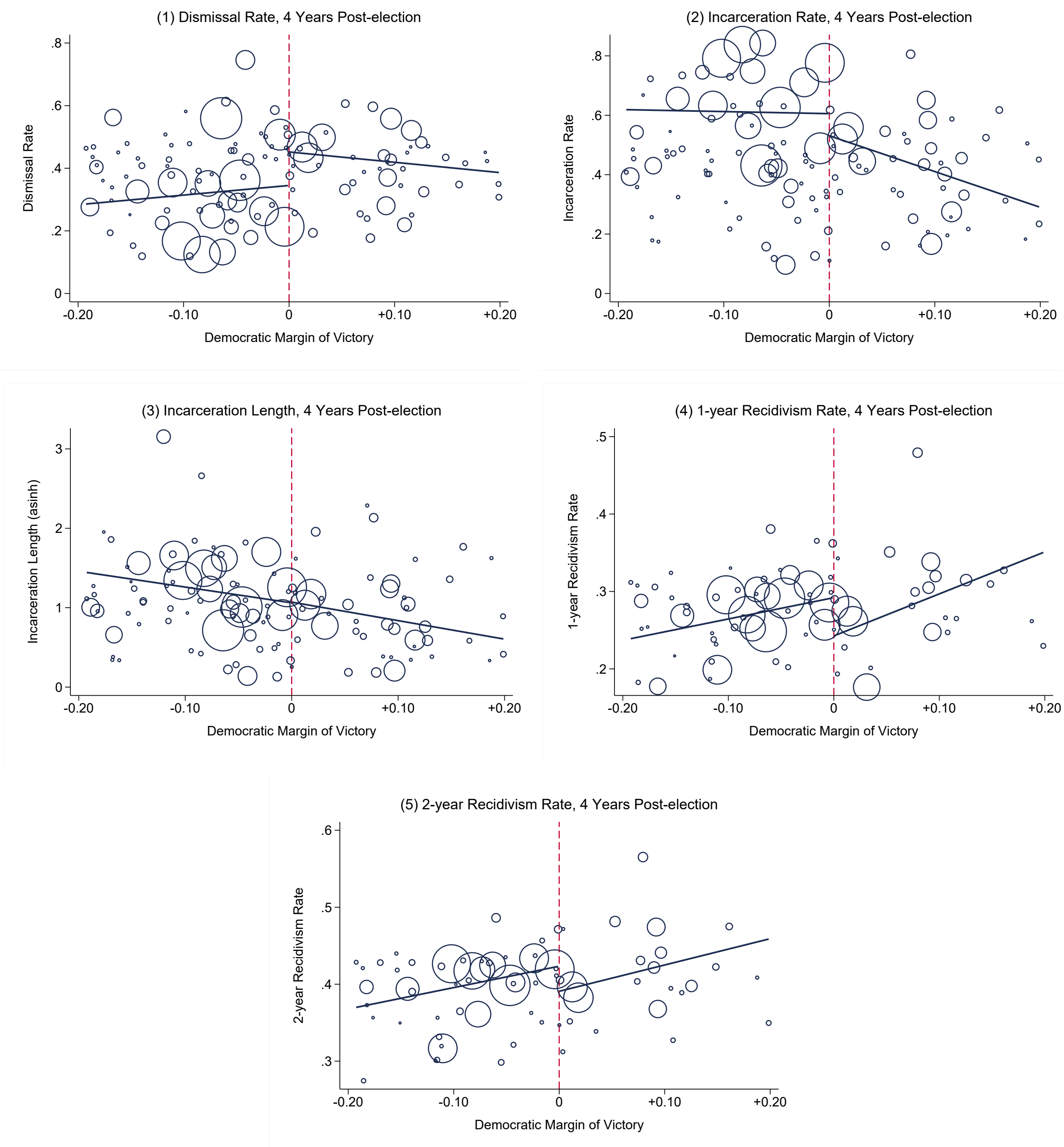


Figure 2.6: Scatter plots show the jurisdiction-level mean outcome in the fourth year after a contested DA election, along with lines of best fit. Each point corresponds to a single jurisdiction-election. Larger points correspond to more populous jurisdictions. For panels 1-3, N=134 jurisdiction-elections. For panels 4 and 5, N=87 jurisdiction-elections.

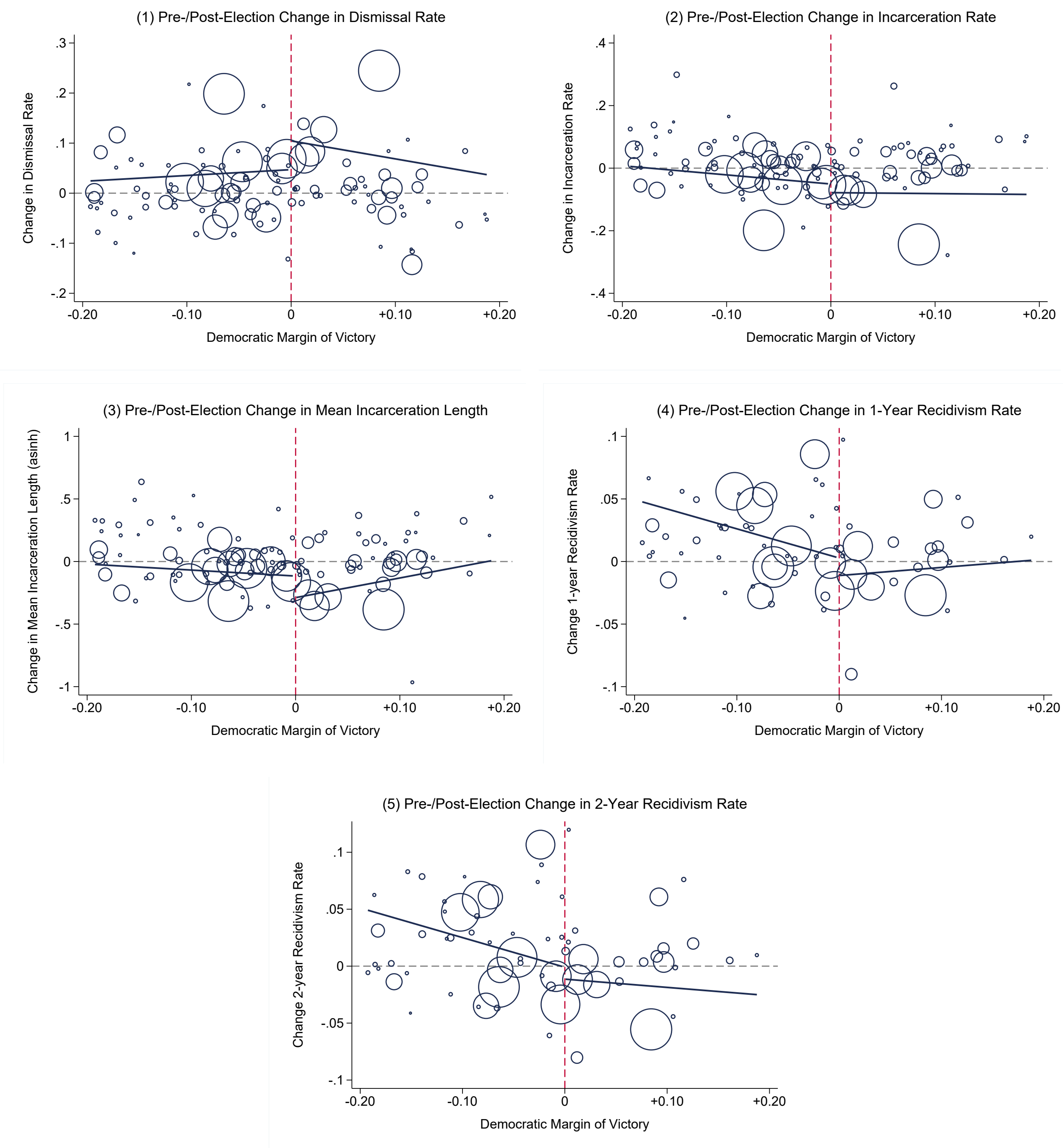
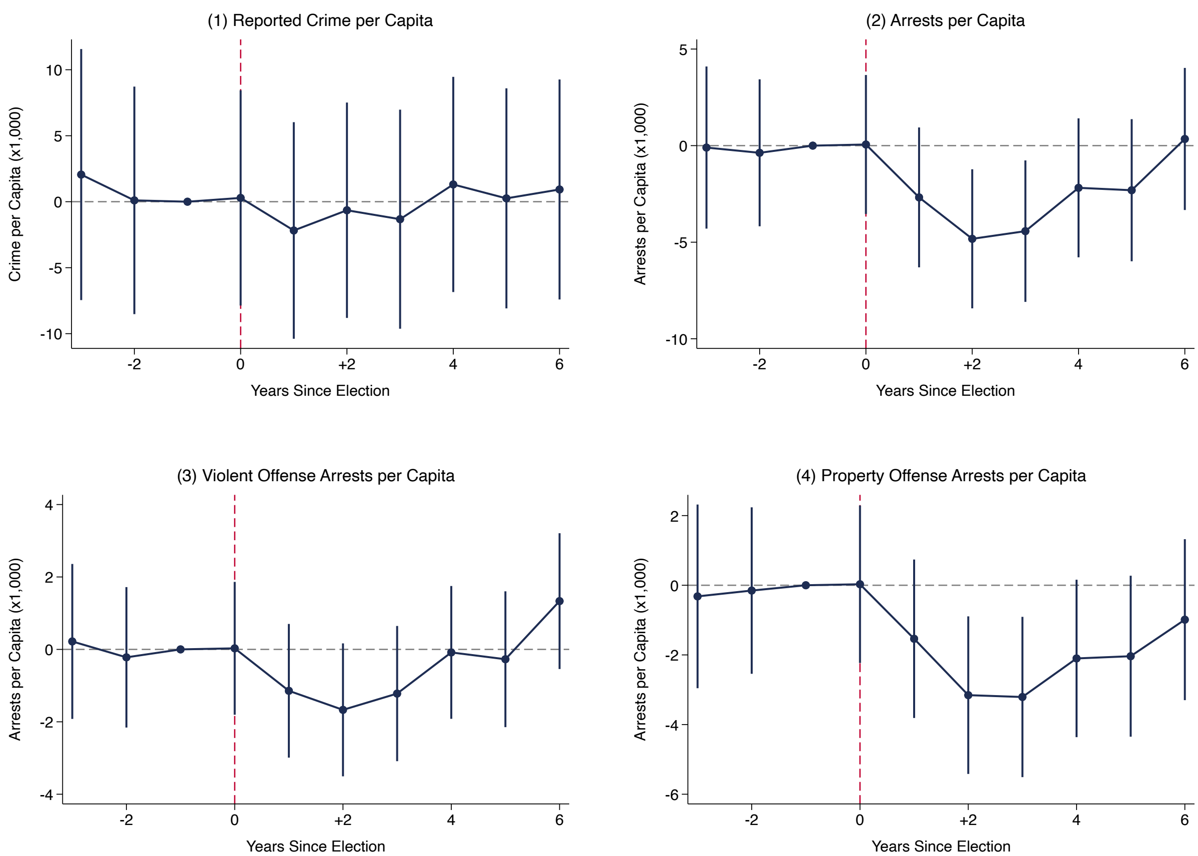


Figure 2.7: Scatter plots show the jurisdiction-level change in mean outcome between the three years leading up to the election and the four years following the election, along with lines of best fit. Each point corresponds to a single jurisdiction-election. Larger points correspond to more populous jurisdictions. For panels 1-3, N=134 jurisdiction-elections. For panels 4 and 5, N=87 jurisdiction-elections.



(1) Reported Crime per Capita
Crime per Capita (x1,000)
Years Since Election
(2) Arrests per Capita
Arrests per Capita (x1,000)
Years Since Election
(3) Violent Offense Arrests per Capita
Arrests per Capita (x1,000)
Years Since Election
(4) Property Offense Arrests per Capita
Arrests per Capita (x1,000)
Years Since Election

Figure 2.8: .

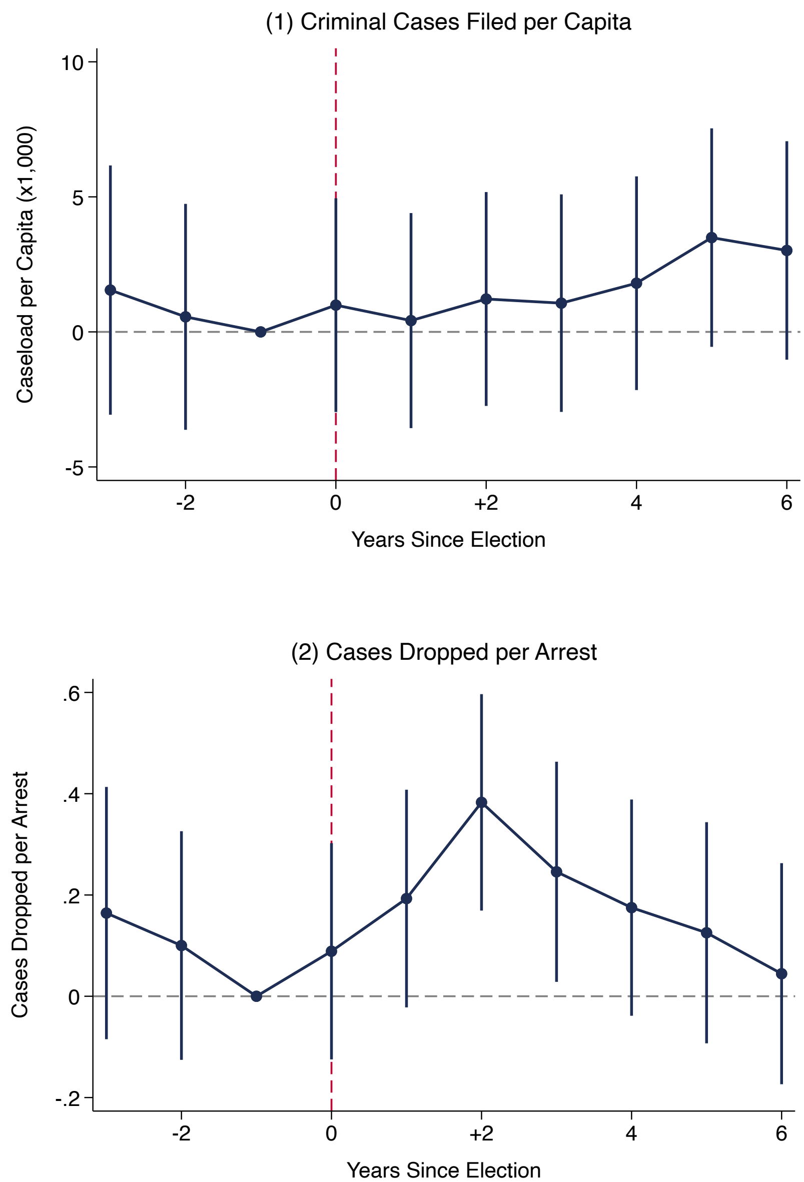

(1) Criminal Cases Filed per Capita
10
5
0
-5
Caseload per Capita (x1,000)
-2
0
+2
4
6
Years Since Election
(2) Cases Dropped per Arrest
.6
.4
.2
0
-.2
Cases Dropped per Arrest
-2
0
+2
4
6
Years Since Election

Figure 2.9: .

Chapter 3

State Interventions into Local Fiscal Stress

3.1 Introduction

In Chapter 1, I discussed how cities alter their spending and public service provision in the face of budgetary pressure, specifically pressure stemming from their workers' defined-benefit pension obligations. In this chapter, I continue to examine local fiscal stress and how state interventions interact with local budgets using oversight and control mechanisms.

State governments have a vested interest in preventing, identifying, and resolving the budgetary challenges faced by their localities. First, local governments are responsible for providing the state's residents' important – and politically visible – services such as public safety, utilities, and transportation, which fiscal stress may impair. Second, poor fiscal conditions within a city could reduce the state's own revenues if economic activity is affected. It can also potentially impact bond ratings for the state and other local governments within the state, harming their ability to borrow (Pew, 2016). Third, while state governments are generally not formally obliged to bail out failing governments, they may find themselves committing financial aid regardless to prevent the consequences of further deterioration.

Thus, states would prefer to detect local fiscal issues early, reducing moral hazard and saving resources they would otherwise need to commit. One prominent example of a city with both budgetary challenges and state intervention is Flint, Michigan. The city encountered multiple financial emergencies in the first two decades of the twenty-first century and the related public health concern due to lead contamination in the poorly-maintained water supply. In this context, the State of Michigan appointed emergency municipal management and sent over $350 million in financial support for the city's recovery.[1]

In order to address financial issues early before they become entrenched, over 20 U.S. states have developed fiscal monitoring programs that regularly evaluate the finances of local governments (Pew, 2016). The exact form of these programs varies. Some are purely informational. Programs, such as New York's Fiscal Stress Monitoring System, act simply to provide up-to-date information on cities, school districts, and other local entities to taxpayers and local officials. While this information is already available in sources such as annual financial reports, the programs present government finances in a format that may be accessible to a wider audience. The information included varies, but programs typically rate local fiscal stress based on operating deficits or surpluses, fund balances, debt levels, pension obligations, and tax revenue growth or decline, among other factors. Other programs include varying types of interventions into local governments by the state, conditional on those local governments' fiscal health. These can take the form of review and approval of budgets, financial audits, and even direct management of local finances.

I primarily focus on whether local budgets are impacted by state fiscal stress monitoring. One potential mechanism for monitoring having a positive impact is by signaling fiscal health to the public with subsequent political action towards improving local budgets. An informational channel requires a few assumptions. First, it requires that, in aggregate, the voting residents of localities prefer fiscal sustainability. Renters, for instance, may instead prefer when cities spend more if they do not account for current and future property tax

[1]The $350 million figure comes from the State of Michigan's website on Flint's crisis: https://www.michigan.gov/flintwater

increases in their rent (i.e. the renter illusion hypothesis). Second, it requires at least some of the resident voters to be (rationally) ignorant of local fiscal health. The oversight programs generally add no new information – local governments already release annual financial reports to the public. The simpler and more widely broadcasted information would need to increase awareness of fiscal issues like deficits, tax declines, and public worker pension costs relative to complicated annual financial reports. Third, local political leaders would need to be disciplined by public perception and voters. Some politicians may face little competition at the local level, or may have support from voting blocs like public-sector workers that benefit from greater compensation and higher spending. Thus, monitoring must produce meaningful and politically effective signals for it to improve local fiscal health. If they do, we would expect localities to cut expenditures or increase revenue collection as their main levers to accomplish this. By using these levers, the localities can reduce deficit spending, build up general fund reserves, and enhance liquidity, all key aspects of fiscal health. On another note, if at least the first two assumptions are true, we would expect housing prices to decline as the market adjusts to negative news about local fiscal health.

With interventions into local finances beyond monitoring, there are additional paths for effects on local budgets. Suppose a city is fiscally stressed, and in response the state government assumes some control over the city's budget to move the city towards financial recovery. For this intervention to change the city's outcomes, we need to assume that the city would not have otherwise instituted similar recovery plans. Perhaps the state has greater ability to implement financial recovery plans (e.g. accounting and budgeting expertise), or perhaps politics in the city is such that it cannot internally commit to a financial recovery plan. If the state intervention does go beyond the local response, then we would expect to see larger cuts to expenditures or increases to revenue collection than what the city would have engaged in on its own.

In order to empirically assess the effectiveness of monitoring programs, I separately consider two state programs. The first state is California, which has a purely informational program. The second state is Ohio, which provides transparent information and includes accompanying direct interventions into fiscally stressed localities. Given the context of each state's program, I employ two different empirical strategies, using a panel-based regression discontinuity method for California and a matched difference-in-differences approach for Ohio. In both cases, I find that the empirical evidence does not allow a conclusion that either form of monitoring system is effective according to the budgetary outcomes I examine, by either raising revenues or lowering expenditures.

The limited existing literature on fiscal stress monitoring programs and municipalities has found mixed impacts, depending on the outcome studied. In the context of New York, Chung and Williams (2021) find in some of their regression discontinuity specifications that cities labeled as mildly stressed by the monitoring system slightly improve their fund balance ratios and deficits. Also in New York, Yang (2022) uses an event study to find that housing prices drop in cities labeled as significantly stressed, suggesting residents do become more informed, but that municipal bond markets do not react much to labelling as financial information is already priced in. In Ohio – a state with a system also studied in my paper and which intervenes beyond just labelling – Thompson (2017) finds large reductions in operating and capital expenditures in response to fiscal stress labels, with an estimated 25% reduction in total expenditures. However this result is based on a generalized difference-in-differences methodology, put into empirical practice with a two-way fixed-effect model. In this context, the methodology fails to create reasonable comparison groups for local governments treated by Ohio's fiscal labelling system. One key issue is that the fiscal stress label, which is the treatment, is largely applied based on deficit spending which is in turn an outcome of interest in the regressions. Therefore, though my sample of locality-year observations produces analogously large and significant results when using a two-way fixed effect model, my preferred model – which matches labelled local governments with comparable units which did not receive a label – does not generate estimates that can confidently support the effectiveness of fiscal stress monitoring programs.

The rest of the chapter is organized in the following manner. First, I discuss my empirical study of California's system, including context, data, methodology, and results. Then I similarly lay out my analysis of Ohio's system. I finally conclude with a few remarks.

3.2 California Fiscal Stress

3.2.1 Background

Beginning in late October 2019, the California State Auditor released information on cities' fiscal stress to the public through an online dashboard called the "local high-risk program". The intention of the online dashboard is to provide state and local officials, residents and interested parties "a data driven view of each city's fiscal health".[2] Using each city's audited financial statements from fiscal year 2017, the local high-risk program assigned scores based on 10 dimensions, which are shown below in Figure 3.1 along with their relative weight. In total across these dimensions, 60% of the assigned score comes from fund reserves, debts, liquidity, and revenue trends. The remaining 40% of the assigned score comes from the status of cities' pensions and other post-employment benefits, demonstrating the Auditor's concern about the role municipal worker's non-wage benefits programs play in their cities' long-term fiscal health. From this fiscal score, the program categorized cities as high, medium, or low risk, a "stoplight" system corresponding to red, yellow, and green. The scores corresponding to each category are shown in Figure 3.2; 18, 240, and 219 cities were assigned each of these designations, respectively.

3.2.2 Data

This paper analyzes the local high-risk program's initial 2019 roll-out. From the Auditor, I gathered each city's fiscal score and their designated level of fiscal risk (low, medium, or

[2] Information on the local high-risk program is available on the Auditor of the State of California's website at: https://www.auditor.ca.gov/local_high_risk/process_methodology

high). I combine this with information on city finances using the Financial Transactions Reports (FTR), a uniform set of financial data collected by the California State Controller. This provides a city-year level dataset including detailed information on both expenditures and revenues. For fiscal characteristics, I focus on the following broader variables, which I will briefly explain. Total revenue for a city consists of two parts: general revenue, which have no restriction on allocation, and functional revenue, which is restricted by law to be used for a specific function. Total expenditure is made up of current expenditures, which are operating costs like staff wages, and non-current expenditures, which are capital investments and long-term debt expenses. Ideally, I would like to calculate a measure of current deficit spending; but there is no "current revenue" category reported by the cities. Instead, I define two deficit measures: one defined as total expenditures less total revenues, and another defined as current expenditures less general revenues. As in the first chapter, where I also use the FTRs, I transform variables to their real per capita equivalents using the California Department of Finance's city population estimates and fiscal year average California CPI. The latter of these is constructed by the Department of Finance using a population-weighted average of the US Bureau of Labor Statistics CPIs for California locations.

I also bring in information on home values within each city. I use Zillow's home value index (ZHVI).[3] This housing market metric is an average of Zillow's estimates of home values within each city, where each home's estimate is constructed based on sale prices, assessed values, and attributes such as square footage and age as well as information from similar nearby homes. The estimates are able to be constructed even for periods in which a home does not sell, and the indices are made for nearly all U.S. places and geographies, providing near universal coverage for my sample of California cities. The specific ZHVI estimate I use is for all homes in the 35th to 65th percentile range, representing the typical home.

The sample includes 465 cities observed over the period of fiscal years 2017 to 2021. This sample size is after I limit it to cities with an average population of 1,000 or more. The

[3]The ZHVI data are available at: https://www.zillow.com/research/data/

average fiscal score is 69.6, and the median California city received a medium risk label. Table 1 presents summary statistics for fiscal year 2019, which is the last year before the roll-out later in that calendar year. The table has means and standard deviations of per capita expenditures and revenues in fiscal year 2020 dollars. The four columns show the full sample as well as the means (standard deviations) for each of the three risk designation groups. Each labeled risk level differs on average across these city finance, population, and home value dimensions. Therefore, attempting to infer the effect of risk designations based on simple pre- versus post- comparisons between each risk category would be problematic due to the existing differences between cities in each category. This fact, along with the local high-risk program's fiscal score-based policy rule, informs my empirical strategy.

3.2.3 Empirical Strategy

In this examination of California's fiscal stress program, I analyze whether fiscal health designations meaningfully impact local finances and decisions on spending and revenue. On the one hand, by increasing awareness of fiscal issues, especially "shrouded" public worker pension debts and costs, cities may be more likely to adjust their finances. On the other hand, the label adds no new information, so another reasonable hypothesis is that they do not affect budgets. The same consideration applies to home values. If signals on city fiscal health are meaningful, we might expect home values to decrease; but if they just repeat existing information then home values would see no change.

I employ a modified regression discontinuity (RD) design. This design mitigates the endogeneity concerns with a naive design comparing outcomes for all cities within each risk designation with one another, since the estimates produced would be contaminated with pre-existing differences across cities with better or worse fiscal health; that is, cities with worse fiscal health would likely have had different budgetary outcomes than those with better fiscal health, even in the absence of California's fiscal stress program. Instead in the RD design, I compare only cities that lie above and below the risk designation thresholds, rather than the full sample. Here, the RD design assumes that cities are unable to manipulate their exact fiscal score and thus cities with fiscal scores immediately surrounding each risk designation threshold are comparable – in terms of finances, demographics, population trends, and other aspects which we cannot observe – in the absence of the designation. Thus, treatment by a risk designation may be considered "as good as random", and that estimation will uncover the impact of fiscal risk designations rather than underlying differences in cities that receive each designation. In this context, manipulation is essentially ruled out, supporting this particular empirical strategy. Fiscal scores were determined from previous years financial data, so for example cities cannot have expanded their general fund balances to increase their fiscal score just enough to improve their risk designation. Still, to test the no-manipulation assumption, I follow the econometric procedures defined in (Calonico et al., 2020) and earlier work. In Figure 3.3, I plot the density of cities along the running variable (fiscal score) centered at the medium and high risk thresholds, where higher scores are better and thus cities right of each threshold have designations of lower risk than those left of each threshold. I find little visual evidence of a jump at either threshold. The increased mass a few points away from left side of the medium risk threshold pulls the local quadratic polynomial up on that side, though we would not expect cities to manipulate their fiscal score in order to receive a worse fiscal rating. Formal tests also detect no jump, with a p-value of 0.98 at the medium risk threshold and 0.79 at the high risk threshold. Note that these tests were conducted using a triangular kernel and are thus weighted towards detecting manipulation very close to the threshold. Uniform weights, which gives observations relatively far from the threshold the same weight as those next to it, yield p-values of 0.40 and 0.58 respectively for the medium and high risk thresholds, similarly rejecting a hypothesis of running variable manipulation at any reasonable confidence level.

Before moving on to my main empirical method, I look for some preliminary visual evidence from a basic RD design. In Figure 3.4, I plot the running variable of fiscal scores against one-year changes in an assortment of fiscal variables. Row 1 has panels with logged current and non-current expenditures, row 2 has logged general and functional revenues, and row 3 has unlogged deficits. In each, I show the two risk designation thresholds as an orange vertical line and a red vertical line. Moving left from the highest fiscal score of 100, crossing the orange line means switching from a low risk to a medium risk label and crossing the red line means switching from a medium risk to a high risk label. For each outcome, there appears to be little change across either threshold. Note, this is looking only at the one year difference, and cities may take longer to shift budgets in response to the treatment, which is not captured in these figures. However, the two-year differences as outcomes – not shown here – also show little visual differences across the threshold.

I use a specification that modifies the RD design by adding a difference-in-difference panel element. That is, I estimate how the fiscal health designation is related to outcomes within cities before and after the fiscal health dashboard was released for cities with fiscal scores close to the threshold. This method allows me to exploit the additional information provided by observing cities across time, providing more precise estimates. This is especially important given the lack of visual evidence in the usual RD specification, so that a result – null or otherwise – has as tight a confidence interval as possible. The panel specification I use, which is similar to the one used in Chapter 2, is:

$$\begin{aligned} Y_{it} = \alpha_0 + \alpha_1 FiscalScore_i + MediumRisk_i(\alpha_2 + \alpha_3 FiscalScore_i) \\ + postFYE2019_{it}(\gamma_1 + \gamma_2 FiscalScore_i + MediumRisk_i(\beta_1 + \beta_2 FiscalScore_i)) \\ + \mu_i + \theta_t + \epsilon_{it} \end{aligned} \tag{3.1}$$

In Equation 3.1, the fiscal score variable is centered around the medium risk threshold, which is a fiscal score of 71.24. As written, this regression specification measures the post-treatment effect of the medium fiscal risk designation relative to the low fiscal risk designation, with β_1 as the coefficient of interest.

Separately, the post-treatment effect of high compared to medium risk can be measured by recentering the fiscal score variable around that threshold (41.76) and rewriting Equation

3.1 with $HighRisk_i$ rather than $MediumRisk_i$. To be clear, these regressions take the form:

$$\begin{aligned} Y_{it} &= \alpha_0 + \alpha_1 FiscalScore_i + HighRisk_i(\alpha_2 + \alpha_3 FiscalScore_i) \\ &\quad + postFYE2019_{it}(\gamma_1 + \gamma_2 FiscalScore_i + HighRisk_i(\beta_1 + \beta_2 FiscalScore_i)) \\ &\quad + \mu_i + \theta_t + \epsilon_{it} \end{aligned} \tag{3.2}$$

Finally, I also modify Equation 3.1 into an event-study approach. This allows me to examine any pre-trends in outcome variables and to see if there are any dynamics in the (short) post-treatment period. Accordingly, the specification becomes:

$$\begin{aligned} Y_{it} &= \alpha_0 + \alpha_1 FiscalScore_i + HighRisk_i(\alpha_2 + \alpha_3 FiscalScore_i) \\ &\quad + \sum_{t=-2}^{t=2}(\kappa_{it} + \rho_t FiscalScore_i + HighRisk_i(\delta_{1t} + \delta_{2t} FiscalScore_i)) \\ &\quad + \mu_i + \theta_t + \epsilon_{it} \end{aligned} \tag{3.3}$$

Each δ_{1t} coefficient measures the difference between the period t and fiscal year 2019 that can be empirically tied to the marginal assignment of a medium fiscal risk label.

I determine the optimal RD bandwidths using the selection procedure from (Calonico et al., 2020). In regressions focusing on the effect of the medium fiscal risk designation, the optimal RD bandwidth only includes cities with fiscal scores within 7.2 points of the threshold. When considering the high risk designation, it only includes cities with scores within 4.8 points of that threshold. There is no overlap between these bandwidths, so the set of cities included in the medium and high risk specifications are different. Lastly, I cluster standard errors in all specifications at the city level to account for serial correlation of outcomes within cities.

3.2.4 Results

Table 2 reports the results from using Equations 3.1 and 3.2. Each row indicates a different fiscal outcome variable. Columns 1 and 2 refer to the unlogged and logged outcomes for the medium risk specification, and columns 3 and 4 are the same for the high risk specification. In this context, I do not find that either the medium risk designation (relative to low risk) or high risk designation (relative to medium risk) have a statistically significant effect on spending or revenue after the dashboard was released. Beyond significance, estimates do not have signs that are consistently in directions to suggest that labels improve fiscal outcomes; for that, we would expect to see increasing total revenue and decreasing current and non-current expenditures.

With ZHVI as the outcome, there also does not seem to be any short-term impacts on home values. Largely due to the limited number of cities designated as high risk, the estimated effects of moving from medium to high risk have particularly wide confidence intervals. The confidence intervals for the medium risk designation are generally less than half the size, which can partly be attributed to the fact that there are over five times the number of city panels.

Figures 3.5, 3.6, 3.7, and 3.8 show the dynamic effects of the medium risk designation as estimated by Equation 3.3. I plot panel RD estimates showing the impact of medium risk rating versus low within 7.2 points or less of the threshold. These estimates are relative to fiscal year 2019, which is the last year before the dashboard was released. Regressions include city and year FEs in the RD panel models. Vertical bars denote 95 percent confidence intervals using robust standard errors clustered at the city level. The layout of the figures are as follows: the first column has estimates with the outcome in real per-capita terms, while the second row has estimates with the outcome in logged real per-capita terms. The outcomes in the two rows in Figure 3.5 are current expenses and non-current expenses (e.g. capital); in Figure 3.6 they are general revenue and tax revenue; and in Figure 3.8 the only row is for the ZHVI measure of home values. Across these figures, there is no evidence that the purely informational risk designations as applied by the California state government have any short-run effects on city finances. For logged ZHVI, there is a 2% estimated decline in home values for cities designated as medium risk as compared to low risk at the threshold for fiscal year 2020, though any potential effect appears to be transient as the fiscal year 2021 estimate is not distinguishable from zero.

I caveat these results on the effects of informational fiscal labelling, since the California fiscal monitoring dashboard rolled out less than half a year before March 2020. It is possible that in a counterfactual world without the impacts of Covid-19, the effects of labelling could be different — if labels were ignored only because cities and their residents were focused on public health measures and tax revenue declines, for instance.

3.3 Ohio Fiscal Stress

3.3.1 Background

Other states besides California have also instituted various forms of fiscal stress monitoring programs. One example is Ohio, which has had a fiscal distress system for localities since 1979, one of the longest-lasting in the United States. Through this system, the Ohio Auditor of State determines which localities are encountering budgetary issues and sends a notice to the fiscally stressed localities. It also informs the public through a detailed press release. Fiscal distress labels are applied according to whether certain conditions exist and their severity. These include budget deficits, low reserve funds, large debts, debt defaults, failure to make payroll, and unauditable financial records. The most common reason a locality is labelled as fiscally distressed is because they ran a deficit in their budget in the most recent fiscal year.

For all fiscal stress labels, localities must submit an acceptable recovery plan to the Auditor and implement it. The strongest, and most common, designation a locality can receive is "fiscal emergency", in which it is placed under a state-determined financial planning and supervision commission. This commission then assumes authority over the locality's finances, with the ability to approve monthly expenditure and revenue plans, approve any new debt issues, and to enter any contracts or agreements it sees fit to restore financial stability.[4] There are two lower designations as well, fiscal caution and fiscal watch, and localities can have their risk designation elevated up to fiscal emergency depending on the status of their fiscal situation and whether they have made sufficient progress toward remedying it. Fiscal emergencies can last indefinitely until it is considered resolved.

There are three types of localities covered by the system: cities, villages, and townships. Ohio's municipal governments, cities and villages, are essentially the same in terms of powers and responsibilities. Villages are automatically classified as cities if the most recent decennial Census shows it as having more than 5,000 residents. Townships are subcounty-level governments that have coverage across the whole state of Ohio. Compared to the municipalities, townships have somewhat more restricted powers and the public services they provide vary between townships, but they are primarily responsible for road maintenance and public services for people who do not reside within a municipality.

3.3.2 Data

I use the publicly available data from the Ohio Auditor of State to determine fiscal stress labels that they have applied. The Auditor lists all instances of labels applied from 1980 onwards. These data include the name of the locality, the date on which the label was applied, the date on which the label was lifted if it has been, the type of fiscal stress label, and some information on why the label was applied. For each calendar year in my sample, Figure 3.9 shows the number of fiscal stress labels applied across each of the three types. Localities are initially labelled throughout the sample, and are not all initially labelled during or immediately after the recession of the early 2000's or the Great Recession. Fiscal stress

[4]Information on fiscal stress labels, their determinants, and what each entails is available at: https://ohioauditor.gov/fiscal/local.html

labels generally last longer than one year; Figure 3.10 instead shows the number of fiscal stress labels active in each year.

I also acquired detailed financial data for each locality. Each year, the Ohio government requires each local government – including other forms not included here such as school districts and counties – to submit an Annual Financial Report (AFR) which are largely standardized across units of a given type. These AFRs include total expenditures and revenues as well as the constituent parts of each, such as property tax within total revenues.

I use intercensal population estimates in order to transform fiscal variables into the per-capita equivalents. These population estimates are produced by the U.S. Census Bureau through the Federal State Cooperative Program for Population Estimates (FSCPE) Program.[5] Lastly, I use the CPI series for Midwestern urban consumers in order to set dollars to their 2020 value.

In all, the combined panel of localities runs from years 2002 through 2021. I present summary statistics by government type in Table 3. Fiscal stress labels are relatively rare within the sample, making up around 4.1% of city-year observations, 2.8% of village-year observations, and only 0.4% of township-year observations. Villages and cities are both types of municipalities, though villages are much smaller than cities: 1,267 average population compared to 27,591 average population. Even so, villages still have comparable per-capita budgets, collecting revenues around three-quarters of what cities do.

3.3.3 Empirical Strategy

Thus, Ohio provides another context in which to determine whether fiscal stress systems are effective. While pure-information systems like California present no evidence of budgetary impacts from risk designations, its possible that state intervention into local budget processes could. The long time-span over which to view the effects of Ohio's program also avoids the potential issues with California's: a short post-period that could be panedemic-

[5] I accessed these estimates through the Missouri Census Data Center: https://mcdc.missouri.edu/

influenced. An approach to identification of Ohio's program effects would be to compare outcomes between localities with and without a designation from the Auditor of State before and after the designation is applied. One way to accomplish estimation in this context with multiple treatment cohorts across time is to use a two-way fixed effects model.[6] This can be written:

$$Y_{it} = \alpha + \beta FiscalStress_{it} + \mu_i + \theta_t + \epsilon_{it} \tag{3.4}$$

where Y_{it} is a budgetary or other outcome, μ_i are locality fixed effects, θ_t are time fixed effects, and ϵ_{it} is an error term clustered by locality. Here, β is intended to recover the impact of fiscal stress labels on the outcomes of interest.

The two-way fixed effects model described above is comparable to the one used in Thompson (2017) to investigate the role of fiscal stress labelling in Ohio. There are a few issues, however. For β to recover the causal effect of the label and any related state intervention on the localities that receive them, then labelled and unlabelled localities would need to have had the same average outcomes, conditional on the included fixed effects or on any covariates. The assumption is problematic given how the Auditor determines which localities receive labels. Selection bias arises since treatment is based on a locality encountering fiscal difficulties, most often in the form of running a deficit in the year prior to treatment. All nonlabelled localities cannot collectively serve as an adequate control group in this scenario. Describing the issue in another way: the treatment is allocated according to the outcome directly or indirectly at baseline, making the effects of treatment more difficult to differentiate from effects of underlying differences when considering the evolution of the outcome. To pull an example from the labor literature, this is similar to how the "Ashenfelter dip" (Ashenfelter, 1978) complicates estimating the treatment effect of job training programs. Just as workers taking up a job training program are more likely to have experienced a recent loss of earnings, so too localities entering a fiscal stress designation are more likely

[6] I am deliberately ignoring the recent advances in difference-in-differences at this point for the purpose of comparing to Thompson (2017).

to have experienced a recent loss of revenues or an increase in expenditures. If localities encountering budgetary difficulties raise revenues and cut expenditures even without state interventions, then β will tend to overstate the impact of these interventions.

There is also an additional issue of selective treatment timing. The Auditor "selects" localities into treatment by observing their most recent fiscal data. Thus, localities labelled by the fiscal stress systems are spread throughout initial treatment cohorts across time. The standard two-way fixed model does not take into account *when* units enter treatment. The model allows for comparisons between treated units and units that were previously treated. Using language from the recent literature on difference-in-differences with varying treatment timing like Callaway and Sant'Anna (2021), those comparisons cloud inferences with "forbidden" or "nonsensical" differences. I will note that since the treatment by a fiscal stress label is relatively rare among localities within the sample, very few of these comparisons will be made and thus this is a lower-order concern.

I use a conditional difference-in-difference approach instead. Before proceeding with specifics, I will provide a brief overview. The approach consists of matching treated units from each treatment cohort with similar untreated units and then using comparisons within matched units to infer the average treatment effect on the treated. This assumes that, conditional on the match on observable characteristics, outcomes for treated and untreated units would follow parallel trends if the groups were both treated or untreated. Another important assumption is common support: that across the observable characteristics in the match, treated and untreated units can both be found. Thus, interpretation of causality relies on the matches being reasonable.

More specifically, I use the procedure described in Dettman et al. (2021), which builds upon the Heckman conditional difference-in-difference estimator developed in Heckman et al. (1997) and Heckman et al. (1998) by allowing for multiple treatment periods. The procedure consists of three main parts. First, the data are pre-processed to create pools of localities that never received a fiscal stress label as potential matches specific to each labelled locality.

Each pool is created based on the nonmissingness of the observable characteristics included in the match in the year before the individual locality's treatment begins. Since each treated unit's pool is specific to when it enters treatment, it takes into account the heterogeneity in timing embedded in this context. Using the pre-processed data, each locality designated by the program as "fiscally stressed" is paired one-to-one with a similar untreated unit from the pool of potential matches. Using nearest-neighbor matching via a distance function, the closest unit is selected; if no suitable match is found, then the treated unit is dropped from the analysis. Treated localities are matched to untreated localities based on per capita budget deficit, the ratio of the budget deficit to revenue, and population one year before the unit is treated, as well as the change in the outcome between two years and one year before treatment. Lastly, an estimate of the average treatment effect for the treated is then taken as the average across all matched units of the difference between the differences before and after treatment began for treated and untreated units in each pair,

3.3.4 Results

First, as a comparison point to Thompson (2017), I estimate Equation 3.4 with logged fiscal variables as outcomes. Table 4 summarizes these, with Panel A considering any fiscal label as the treatment and Panel B only including the fiscal emergency label as the treatment. The results are qualitatively similar to Thompson's main estimates.[7] Specifically, the two-way fixed effect model estimates that fiscal stress labels are associated with reductions of total expenditures by 13.8%, with a larger percentage reduction of capital expenditures by 29.0%. If interpreted as causal, this would suggest that there is strong evidence that Ohio's fiscal stress labels are effective tools to resolve fiscal distress.

Now, I turn to the conditional difference-in-difference method. In this method I match localities that were labelled as fiscally distressed by the Ohio Auditor of State to suitable

[7] They are not the same for a few reasons. The main factor is that I cover a different and longer time period, with the years 2002 to 2021 versus Thompson's time period of 1999 to 2012. I also do not include the local economic characteristics as covariates, though it is unlikely these shift estimates much after the inclusion of two-way fixed effects.

controls and then employ difference-in-differences across the matched units.

Table 5 contains the estimates from this method. Column 1 contains each of the difference-in-difference estimates, with the robust standard error in parentheses below it. Columns 2 and 3 contain the component differences which make up the DID, with the temporal change in the outcome for the treated and for the controls, respectively. Column 4 lists the total number of observations. Like Table 4, Table 5 is split into two panels, Panel A for estimates including all types of fiscal stress labels and Panel B for estimates including only the strongest and most common label of fiscal emergency. Lastly, outcomes are log total revenue and log total expenditure measured one, two, and three years after the treated locality is first labelled as fiscally stressed.[8]

From the estimates, there is little evidence that the fiscal emergency label or any stress designation reduces local deficits by either raising revenues or lowering expenditures. On the other hand, we also cannot conclude that fiscal stress labels and accompanying state interventions do not improve local fiscal conditions. Contributing to this lack of statistical significance is that there are simply too few localities that have been treated by fiscal labels, leading to particularly large standard errors. Thus, while true effects of the size reported in Table 4 and in Thompson (2017) are within reasonable confidence intervals of the conditional DID results, we cannot measure them when using reasonable control groups to account for the selection issues inherent to Ohio's fiscal stress labelling system.

3.4 Conclusion

There is a small literature that considers the role that interventions into local finances by state governments – taking forms such as budgetary monitoring, fiscal labelling, and measures of control – plays in local outcomes. This work contributes to that literature by

[8] I do not estimate effects on capital expenditures using the conditional difference-in-difference method – even though estimates were large and significant in the previous table – as there are about half as many matches as with total revenue and expenditures. This is because there are about 10,000 fewer locality-year observations of capital expenditures; the conditional difference-in-difference method is sensitive to missingness as it cannot match on previous values of the outcome when missing.

suggesting that even in the setting of Ohio, which has laws that allow for direct interventions into local finances that are less rare than most other states, there are few conclusions that can be made empirically about these interventions. Researchers looking into this area must consider the context of how each state decides which locality receives fiscal labels or other interventions when deciding their research methodology.

3.5 Figures

Figure 3.1: Points Possible By Financial Indicators

FINANCIAL INDICATORS	POINTS POSSIBLE
1. General Fund Reserves	30
2. Debt Burden	15
3. Liquidity	10
4. Revenue Trends	5
5. Pension Obligations	10
6. Pension Funding	5
7. Pension Costs	5
8. Future Pension Costs	5
9. OPEB Obligations	10
10. OPEB Funding	5
Maximum Score Possible	**100**

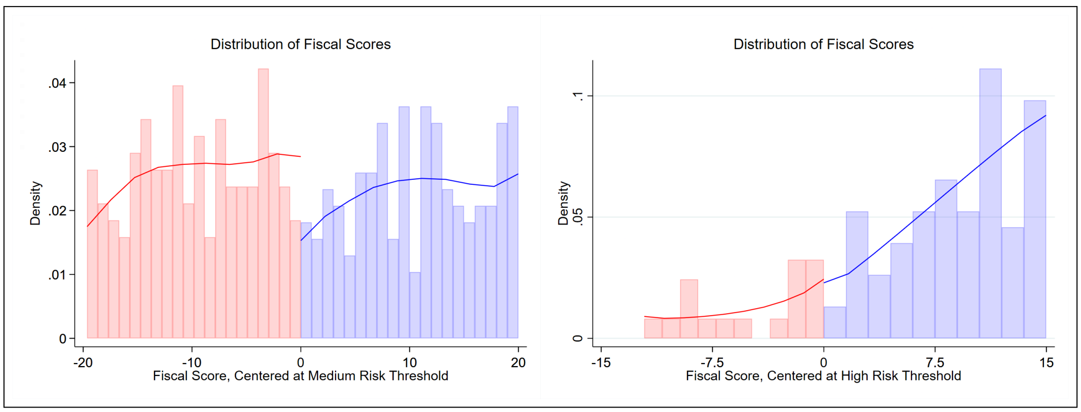
Source: Local Government High-Risk Dashboard

Figure 3.2: Methodology for Determining Which Cities Are High Risk

RISK DESIGNATION	RANGE OF POINTS ASSIGNED	DESCRIPTION
HIGH RISK	0 to 41.76	This designation means that a city has significant risk of experiencing fiscal distress.
MODERATE RISK	41.77 to 71.23	This designation means that a city has some risk of experiencing fiscal distress.
LOW RISK	71.24 to 100	This designation indicates that a city has low risk of experiencing fiscal distress.

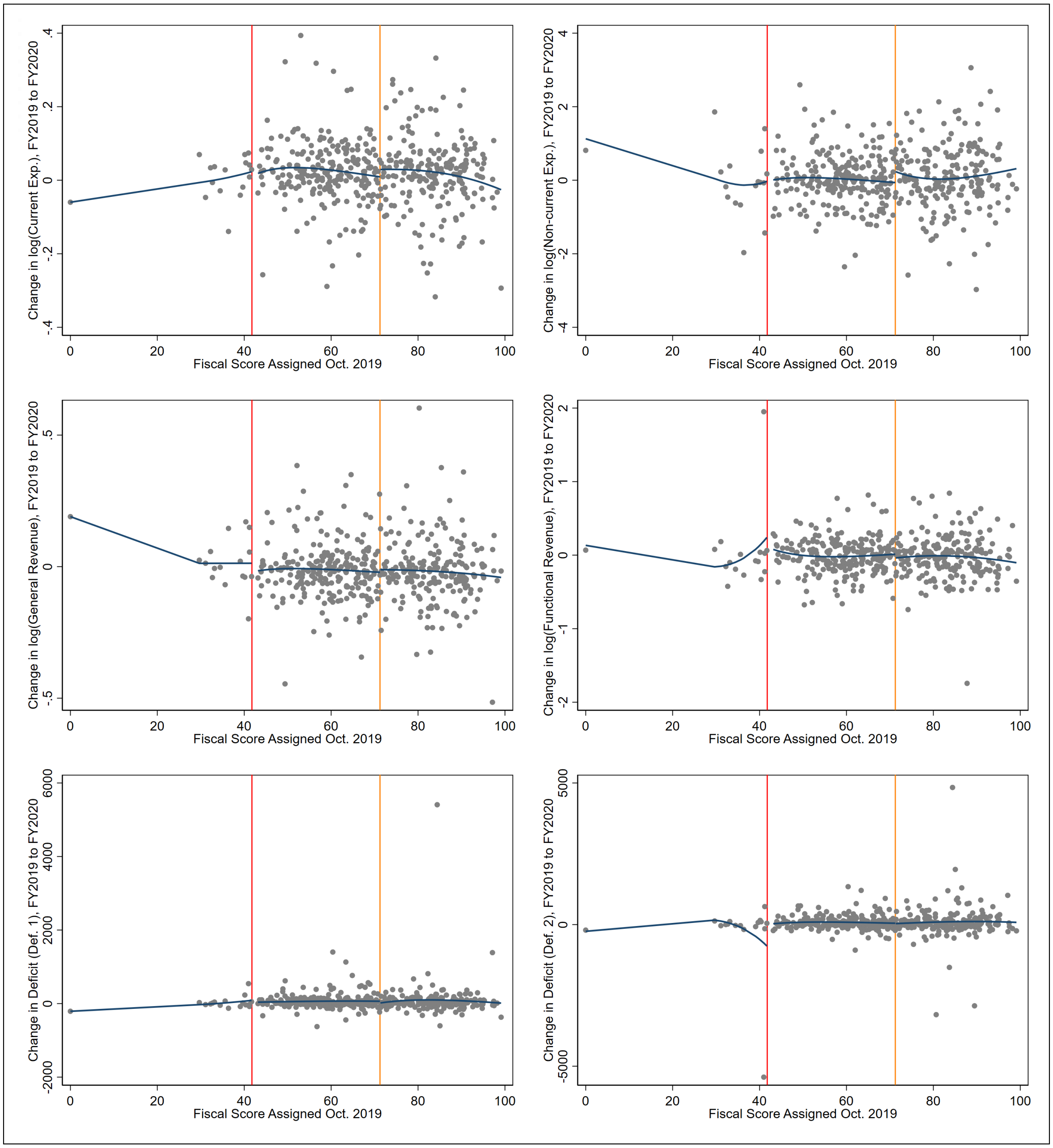
Source: Local Government High-Risk Dashboard

Figure 3.3: Densities for Assessing Manipulation



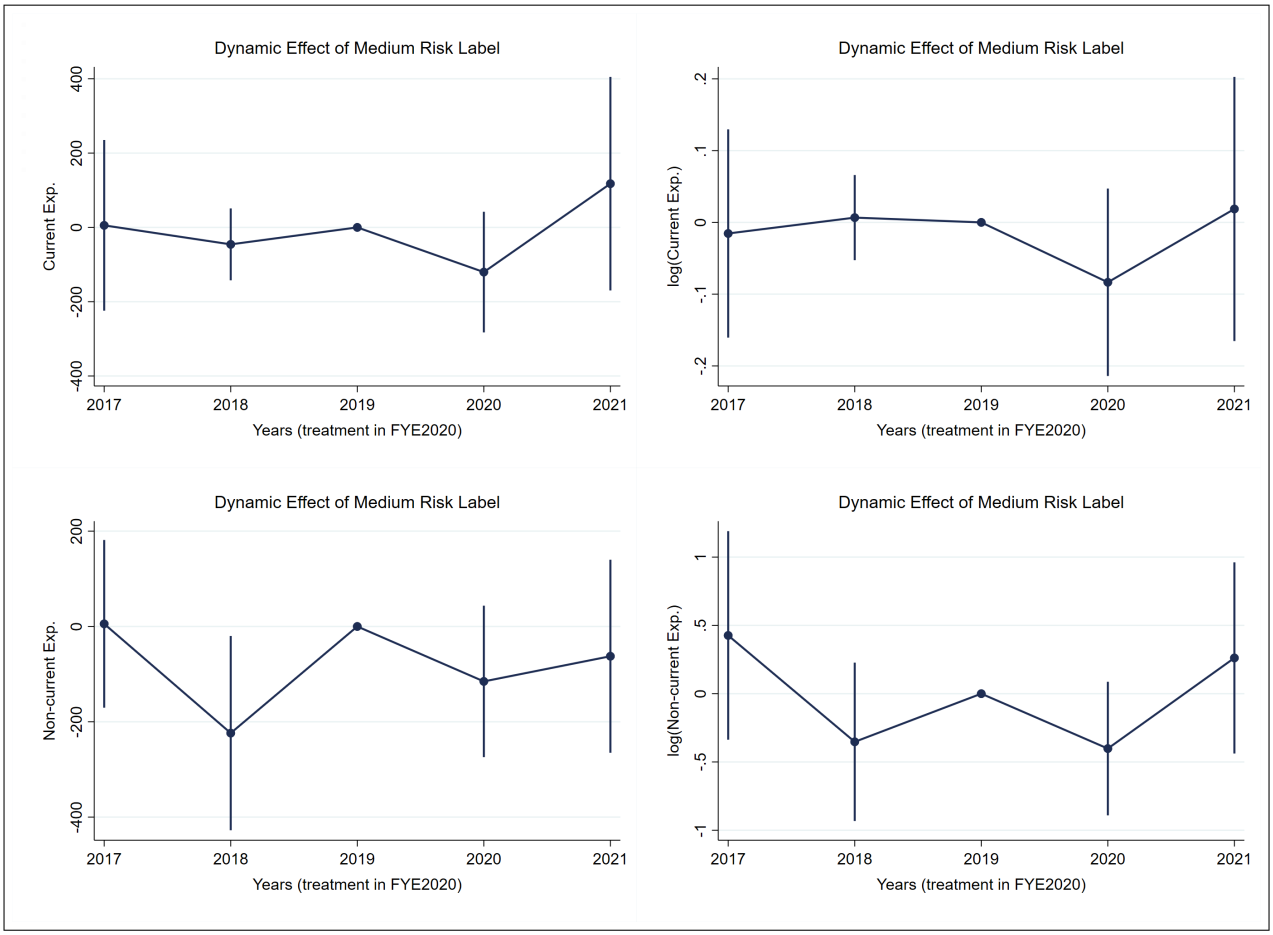
Panel 1 shows the binned density of fiscal scores 20 points above (low risk) and below (medium risk) of the threshold for a California city receiving a medium risk label. Panel 2 shows the binned density of fiscal scores 15 points above (medium risk) and below (high risk) of the threshold for a California city receiving a high risk label. Both panels include local quadratic polynomials which are discontinuous at the threshold. Source: Financial data from the Local Government High-Risk Dashboard

Figure 3.4: Change in Fiscal Variables from FYE 2019 to FYE 2020



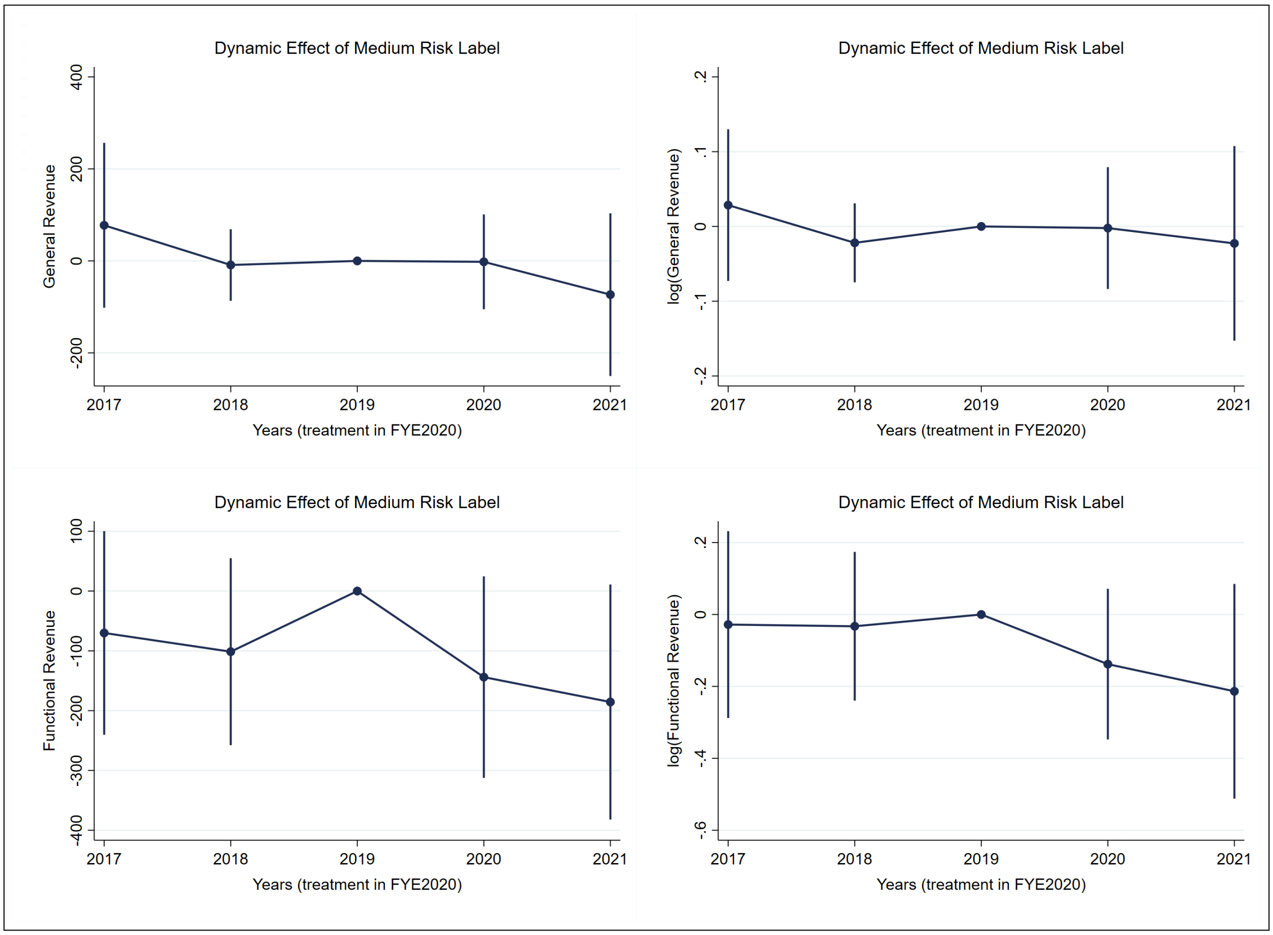
Row 1 has panels with logged current and non-current expenditures, row 2 has logged general and functional revenues, and row 3 has unlogged deficits. Source: Financial data from the Local Government High-Risk Dashboard combined with unified financial reports from the California State Controller.

Figure 3.5: Dynamic Panel Regression Discontinuity Results: Expenditure



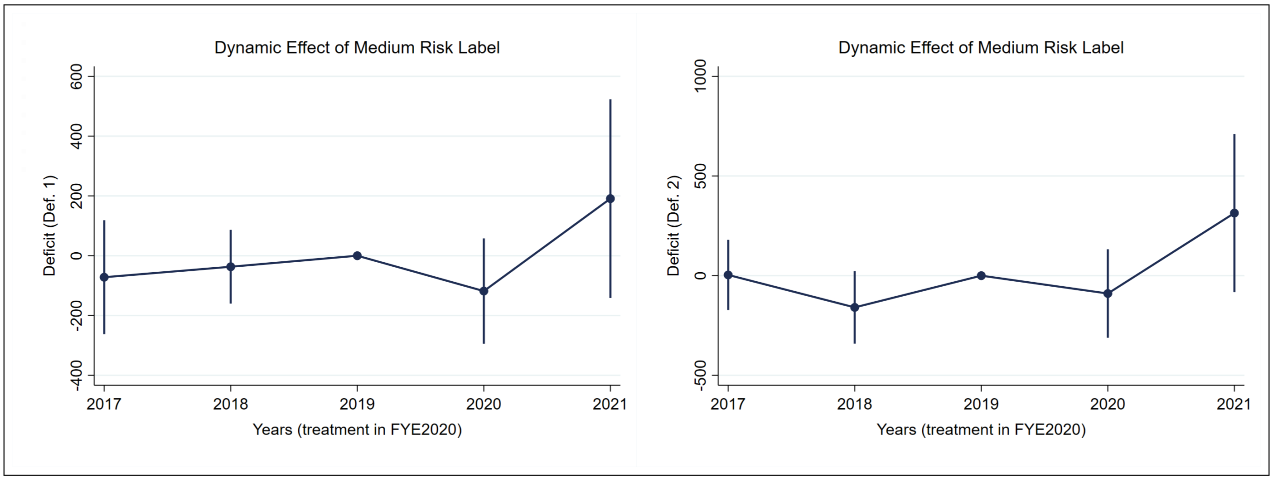
Graphs show the results of Equation 3 for each of the two expenditure outcomes I study. Unlogged outcomes appear in the first column, and logged outcomes in the second column. Each year is compared to fiscal year 2019, which is the last year before the fiscal labelling system was rolled out. Vertical bars denote 95 percent confidence intervals using robust standard errors clustered at the city level. Source: Financial data from the Local Government High-Risk Dashboard combined with unified financial reports from the California State Controller.

Figure 3.6: Dynamic Panel Regression Discontinuity Results: Revenue



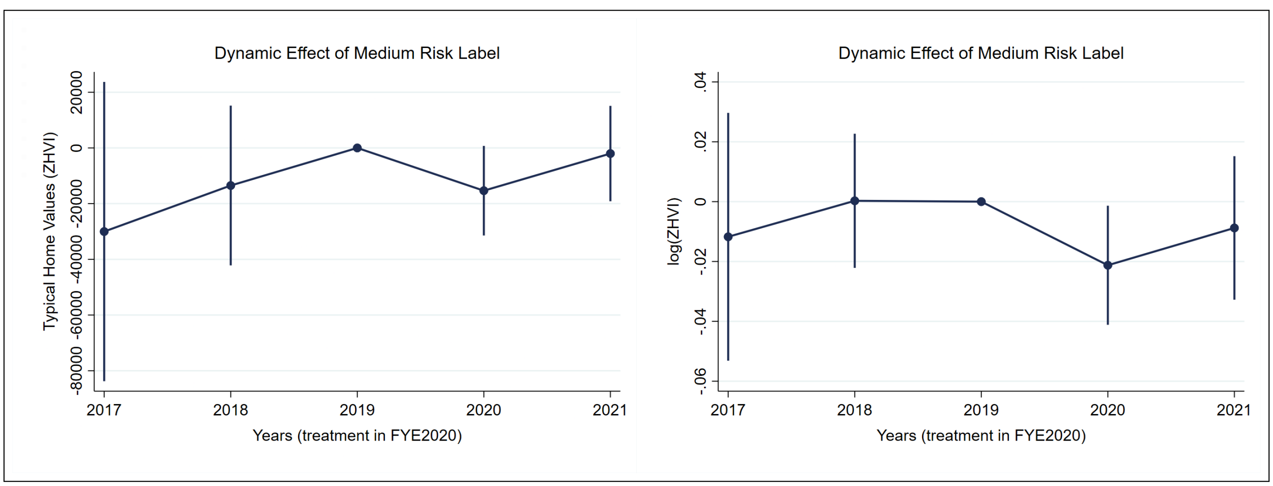
Graphs show the results of Equation 3 for each of the two revenue outcomes I study. Unlogged outcomes appear in the first column, and logged outcomes in the second column. Each year is compared to fiscal year 2019, which is the last year before the fiscal labelling system was rolled out. Vertical bars denote 95 percent confidence intervals using robust standard errors clustered at the city level. Source: Financial data from the Local Government High-Risk Dashboard combined with unified financial reports from the California State Controller.

Figure 3.7: Dynamic Panel Regression Discontinuity Results: Per-Capita Deficits



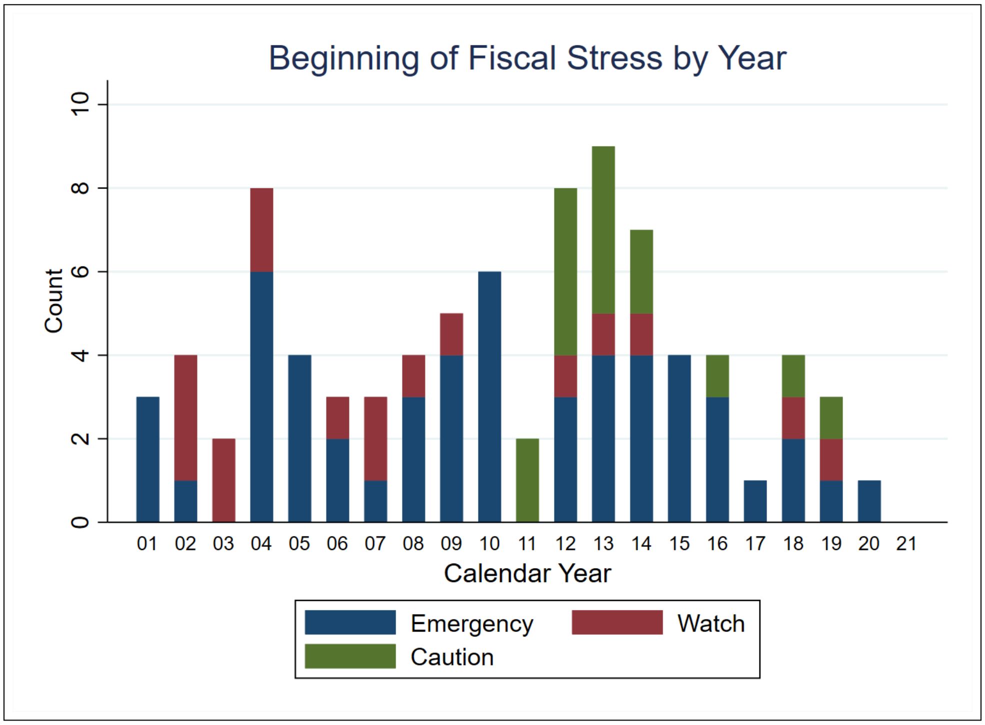
Graphs show the results of Equation 3 for each of the two versions of the per-capita deficit measures used. The first is defined as total expenditures less total revenues, and the second as current expenditures less general revenues. Each year is compared to fiscal year 2019, which is the last year before the fiscal labelling system was rolled out. Vertical bars denote 95 percent confidence intervals using robust standard errors clustered at the city level. Source: Financial data from the Local Government High-Risk Dashboard combined with unified financial reports from the California State Controller.

Figure 3.8: Dynamic Panel Regression Discontinuity Results: Home Values



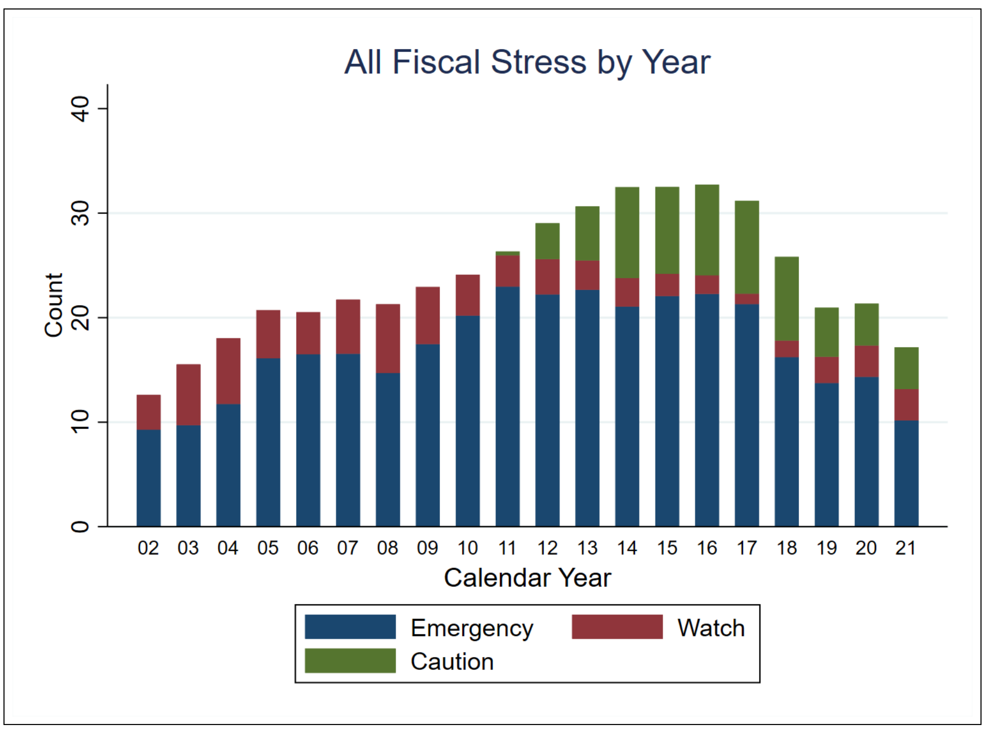
Graphs show the results of Equation 3 where home values are the outcome, as measured by the Zillow home value index (ZHVI) for all homes in the 35th to 65th percentile range, smoothed and seasonally adjusted. Unlogged outcomes appear in the first column, and logged outcomes in the second column. Each year is compared to fiscal year 2019, which is the last year before the fiscal labelling system was rolled out. Vertical bars denote 95 percent confidence intervals using robust standard errors clustered at the city level. Source: Financial data from the Local Government High-Risk Dashboard combined with unified financial reports from the California State Controller.

Figure 3.9: Initial Fiscal Stress Designations by Year



Source: Fiscal stress designations by the Ohio Auditor of State.

Figure 3.10: Ongoing Fiscal Stress Designations in each Year



Source: Fiscal stress designations by the Ohio Auditor of State.

3.6 Tables

Table 1: Summary Statistics – California

	(1) Full Sample	(2) Low Risk	(3) Medium Risk	(4) High Risk
Total Revenue	1,585 (1,570)	1,710 (2,032)	1,492 (1,044)	1,328 (799)
General Revenue	1,042 (1,301)	1,149 (1,713)	965 (818)	790 (507)
Functional Revenue	542 (383)	560 (450)	527 (315)	538 (329)
Total Expenditure	1,501 (1,590)	1,611 (2,129)	1,418 (929)	1,281 (756)
Current Expenditure	1,230 (1,324)	1,298 (1,767)	1,182 (790)	1,064 (623)
Non-current Expenditure	270 (374)	313 (498)	236 (222)	217 (183)
Deficit (Def. 1)	188 (285)	148 (309)	217 (259)	274 (263)
Deficit (Def. 2)	-84 (313)	-99 (391)	-74 (234)	-47 (145)
Population	70,525 (212,572)	38,774 (43,655)	99,561 (291,689)	63,401 (98,939)
ZHVI	744,326 (676,129)	892,667 (877,190)	635,516 (420,350)	504,053 (232,280)
N:	*465*	*212*	*236*	*17*

The data describe the fiscal characteristics and population of Californian cities in fiscal year 2019, limited to those with a population above 1,000. Financial variables are reported in per-capita 2020 dollars. There are two deficit variables: one defined as total expenditures less total revenues, and another defined as current expenditures less general revenues. ZHVI is Zillow's home value index for all homes in the 35th to 65th percentile range, smoothed and seasonally adjusted. Column 1 reports the mean across all cities. Columns 2, 3, and 4 report the means for cities that were designated as low, medium, and high fiscal risk respectively by the State Auditor of California. Standard deviations appear in parentheses.

Table 2: Results: Regression Discontinuity DID (California)

	(1) Med. Risk(log)	(2) Med. Risk(unlogged)	(3) High Risk (log)	(4) High Risk (unlogged)
Current Exp.	-0.029	12.075	0.075	81.809
	(0.061)	(97.684)	(0.167)	(298.920)
Non-current Exp.	-0.093	-16.122	0.661	183.453
	(0.242)	(71.151)	(0.649)	(181.697)
General Rev.	-0.015	-60.520	0.102	29.610
	(0.053)	(78.506)	(0.119)	(106.103)
Functional Rev.	-0.156	-107.575	0.445	801.658
	(0.104)	(68.483)	(0.390)	(761.603)
Deficit (Ver. 1)		72.595		52.199
		(104.568)		(352.606)
Deficit (Ver. 2)		115.750		124.668
		(106.095)		(351.003)
ZHVI	-0.011	4615.992	-0.039	-21369.514
	(0.014)	(13878.183)	(0.083)	(35119.785)
N:	*640*	*640*	*115*	*115*

Notes: Estimates come from regressions including the triple interaction of a post-FYE2019 indicator, the relevant indicator for the assigned risk designation, and the fiscal score, where the estimate shown is the interaction between the risk designation and post-FYE2019 indicator. Cities with fiscal scores below 71.3 were assigned Medium risk, and cities with fiscal scores below 42 were assigned High risk. Bandwidths, as determined by optimum bandwidth selection procedures, are 7.2 fiscal points and 4.8 points around the cutoff, respectively. Expenditure variables are presented as per capita, 2020 dollars, both logged and unlogged. Medium and High risk designations were determined by the California State Auditor in October 2019. ZHVI is Zillow's home value index for all homes in the 35th to 65th percentile range, smoothed and seasonally adjusted. Standard errors are given in parentheses, and are clustered by city. Sample sizes are in brackets. The coefficients' significance levels are symbolized as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Summary Statistics – Ohio

	(1) Cities	(2) Villages	(3) Townships
I. Locality-Year Fiscal Stress Counts			
Emergency	110	236	42
Watch	54	34	3
Caution	17	23	38
II. Fiscal Characteristics			
Total Revenue	1,274	898	305
	(695)	(2,071)	(567)
Total Tax Revenue	801	369	130
	(557)	(816)	(383)
Total Expenditure	1,324	877	288
	(787)	(1629)	(523)
Total Capital Exp.	192	155	31
	(254)	(528)	(102)
Population	27,591	1,267	4,262
	(65,269)	(1,285)	(6,978)
N:	*4,397*	*10,436*	*23,576*

Panel I includes counts of the number of locality-year observations for each of the types of fiscal stress labels. Panel II describes the fiscal characteristics and population of Ohio localities. Fiscal variables are reported in per-capita 2020 dollars. Columns 1, 2 and 3 report the mean for cities, villages, and townships, respectively. Standard deviations appear in parentheses.

Table 4: Results: Basic DID Estimates

	(1) Total Revenue	(2) Total Tax Rev.	(3) Total Exp.	(4) Capital Exp.
Panel A: Any Fiscal Stress Label				
Fiscal Stress	-0.051	-0.017	-0.148***	-0.343**
	(0.030)	(0.041)	(0.033)	(0.146)
R-squared:	0.883	0.885	0.852	0.519
N:	*37,631*	*37,502*	*37,624*	*27,891*
Panel B: Fiscal Emergencies Only				
Fiscal Emergency	-0.047	0.002	-0.164***	-0.374*
	(0.038)	(0.051)	(0.041)	(0.192)
R-squared:	0.883	0.885	0.852	0.519
N:	*37,631*	*37,502*	*37,624*	*27,891*

Notes: Estimates come from the two-way fixed effect model written as Equation 3.4. Outcomes are the natural logs of the per capita variables listed. Panel A includes all fiscal stress labels (emergency, watch, caution), whereas Panel B only considers the fiscal emergency label. Standard errors are clustered by locality. The coefficients' significance levels are symbolized as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Results: Conditional DID Estimates

	(1) DID Est.	(2) Mean ΔTreatment	(3) Mean ΔControl	(4) Obs.
Panel A: Any Fiscal Stress Label				
Total Rev., 1 year	0.123	0.052	-0.071	38
	(0.100)			
Total Rev., 2 year	0.058	0.154	0.097	41
	(0.119)			
Total Rev., 3 year	-0.035	0.172	0.206	35
	(0.185)			
Total Exp., 1 year	0.063	0.041	-0.022	39
	(0.132)			
Total Exp., 2 year	0.087	0.095	0.008	43
	(0.165)			
Total Exp., 3 year	0.072	0.067	-0.005	35
	(0.181)			
Panel B: Fiscal Emergencies Only				
Total Rev., 1 year	0.070	0.055	-0.014	34
	(0.146)			
Total Rev., 2 year	0.083	0.197	0.114	37
	(0.146)			
Total Rev., 3 year	0.046	0.222	0.176	30
	(0.178)			
Total Exp., 1 year	-0.018	0.047	0.065	36
	(0.103)			
Total Exp., 2 year	0.040	0.117	0.077	40
	(0.137)			
Total Exp., 3 year	0.029	0.073	0.044	32
	(0.145)			

Notes: Estimates come from the conditional difference-in-difference procedure from Dettman et al. (2021). Treated localities are matched to untreated localities based on per capita budget deficit, the ratio of the budget deficit to revenue, and population one year before the unit is treated, as well as the change in the outcome between two years and one year before treatment. Outcomes are the natural logs of the per capita variables listed. Panel A includes all fiscal stress labels (emergency, watch, caution), whereas Panel B only considers the fiscal emergency label. Abadie Imbens (2006, 2011) robust standard errors are used. The coefficients' significance levels are symbolized as: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 4

Bibliography

Agan, Amanda and Sonja Starr, "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment," *The Quarterly Journal of Economics*, 2017, *133* (1), 191–235.

_ **, Matthew Freedman, and Emily Owens**, "Is Your Lawyer a Lemon? Incentives and Selection in the Public Provision of Criminal Defense," *The Review of Economics and Statistics*, 2021, *103* (2), 294–309.

Agan, Amanda Y, Jennifer L Doleac, and Anna Harvey, "Misdemeanor Prosecution," 2021, *NBER Working Paper No. 28600.*

_ **,** _ **, and** _ , "Prosecutorial Reform and Local Crime Rates," 2021, *SSRN Working Paper Series.*

Alschuler, Albert W., "The Prosecutor's Role in Plea Bargaining," *University of Chicago Law Review*, 1968, *36* (1).

Andonov, Aleksandar and Joshua D Rauh, "The Return Expectations of Public Pension Funds," *The Review of Financial Studies*, 11 2021, *35* (8), 3777–3822.

Anzia, Sarah F., "Pensions in the Trenches: How Pension Spending Is Affecting U.S. Local Government," *Urban Affairs Review*, 2022, *58* (1), 3–32.

Arora, Ashna, "Too Tough on Crime? The Impact of Prosecutor Politics on Incarceration," 2019, *Working Paper.*

Ashenfelter, Orley, "Estimating the Effect of Training Programs on Earnings," *Review of Economics and Statistics*, 1978, *60* (1), 47–57.

Augustine, Elsa, Johanna Lacoe, Steven Raphael, and Alissa Skog, "The Impact of Felony Diversion in San Francisco," *Journal of Policy Analysis and Management*, 2022.

Autor, David, Christopher J. Palmer, and Parag Pathak, "Gentrification and the Amenity Value of Crime Reductions: Evidence from Rent Deregulation," NBER Working Papers 23914, National Bureau of Economic Research, Inc 2017.

Bagchi, Sutirtha, "The effects of political competition on the generosity of public-sector pension plans," *Journal of Economic Behavior & Organization*, 2019, *164*, 439–468.

Baicker, Katherine, "The Budgetary Repercussions Of Capital Convictions," *The B.E. Journal of Economic Analysis & Policy*, August 2004, *4* (1), 1–28.

Beach, Brian and Daniel B. Jones, "Gridlock: Ethnic Diversity in Government and the Provision of Public Goods," *American Economic Journal: Economic Policy*, 2017, *9* (1), 112–36.

Besley, Timothy J. and Stephen Coate, "An Economic Model of Representative Democracy," *Quarterly Journal of Economics*, 1997, *112* (1), 85–114.

Bibas, Stephanos, "Plea Bargaining outside the Shadow of Trial," *Harvard Law Review*, 2004, *117* (8), 2463–2547.

Bouton, Laurent, Alessandro Lizzeri, and Nicola Persico, "The Political Economy of Debt and Entitlements," *The Review of Economic Studies*, 01 2020, *87* (6), 2568–2599.

Bowers, Josh, “Legal Guilt, Normative Innocence, and The Equitable Decision Not to Prosecute,” *Columbia Law Review*, 2010, *110* (7).

Brown, Jeffrey and Richard F. Dye, “Illinois Pensions in a Fiscal Context: A (Basket) Case Study,” NBER Working Papers 21293, National Bureau of Economic Research, Inc 2015.

_ , Joshua M. Pollet, and Scott Weisbenner, “The In-State Equity Bias of State Pension Plans,” NBER Working Papers 21020, National Bureau of Economic Research, Inc 2015.

Callaway, Brantly and Pedro H. C. Sant’Anna, “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, 2021, *225* (2), 200–230.

Calonico, Sebastian, Matias Cattaneo, and Max Farrell, “Optimal Bandwidth Choice for Robust Bias-corrected Inference in Regression Discontinuity Designs,” *The Econometrics Journal*, 2020, *23* (2), 192–210.

CalPERS, “2015-16 Comprehensive Annual Financial Report,” Technical Report 2016.

Chung, Il Hwan and Daniel Williams, “Local governments’ responses to the fiscal stress label: the case of New York,” *Local Government Studies*, 2021, *47* (5), 808–835.

Cohen, Alma and Crystal S. Yang, “Judicial Politics and Sentencing Decisions,” *American Economic Journal: Economic Policy*, 2019, *11* (1), 160–91.

Cromwell, Erich and Keith Ihlanfeldt, “Local Government Responses to Exogenous Shocks in Revenue Sources: Evidence From Florida,” *National Tax Journal*, 2015, *68* (2), 339–376.

Dettman, Eva, Alexander Giebler, and Antje Weyh, “Flexpaneldid: A Stata toolbox for causal analysis with varying treatment time and duration,” *Stata Journal (under review*, 2021.

Dippel, Christian, "Political Parties Do Matter in US Cities... for Their Unfunded Pensions," *American Economic Journal: Economic Policy*, 2022, *14* (3), 33–54.

_ **and Michael Poyker**, "How Common are Electoral Cycles in Criminal Sentencing?," 2019, *NBER Working Paper No. 25716.*

Dobbie, Will, Jacob Goldin, and Crystal S. Yang, "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges," *American Economic Review*, 2018, *108* (2), 201–40.

Drazen, Allan and Marcela Eslava, "Electoral manipulation via voter-friendly spending: Theory and evidence," *Journal of Development Economics*, 2010, *92* (1), 39–52.

Feigenberg, Benjamin and Conrad Miller, "Racial Divisions and Criminal Justice: Evidence from Southern State Courts," *American Economic Journal: Economic Policy*, 2021, *13* (2), 207–40.

Feler, Leo and Mine Z. Senses, "Trade Shocks and the Provision of Local Public Goods," *American Economic Journal: Economic Policy*, 2017, *9* (4), 101–43.

Ferreira, Fernando and Joseph Gyourko, "Do Political Parties Matter? Evidence from U.S. Cities," *The Quarterly Journal of Economics*, 2009, *124* (1), 399–422.

Fischer, Brett, "No Spending Without Representation: School Boards and the Racial Gap in Education Finance," *American Economic Journal: Economic Policy*, 2022.

Glaeser, Edward L., "Chapter 4 - Urban Public Finance," in Alan J. Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, eds., *Handbook of Public Economics*, Vol. 5, Elsevier, 2013, pp. 195–256.

_ **and Giacomo A.M. Ponzetto**, "Shrouded costs of government: The political economy of state and local public pensions," *Journal of Public Economics*, 2014, *116*, 89–105.

Gramlich, John, "Voters' perceptions of crime continue to conflict with reality," 2016.

_ , "U.S. public divided over whether people convicted of crimes spend too much or too little time in prison," 2021.

Heckman, James, Hidehiko Ichimura, and Petra E. Todd, "Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, 1997, *64* (4), 605–654.

_ , _ , Jeffrey Smith, and Petra Todd, "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 1998, *66* (5), 1017–1098.

Jerch, Rhiannon, Matthew Kahn, and Gary C. Lin, "Local Public Finance Dynamics and Hurricane Shocks," NBER Working Papers 28050, National Bureau of Economic Research, Inc 2020.

Kaplan, Jacob, "Uniform Crime Reporting Program Data: Law Enforcement Officers Killed and Assaulted (LEOKA) 1960-2020 [dataset]," `https://doi.org/10.3886/E102180V11` 2021. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Kelley, Dashle, "The political economy of unfunded public pension liabilities," *Public Choice*, 2014, *158* (1), 21–38.

Kling, Jeffrey R., "Incarceration Length, Employment, and Earnings," *American Economic Review*, 2006, *96* (3), 863–876.

Komarek, Tim and Gary A. Wagner, "Local Fiscal Adjustments From Depopulation: Evidence From The Post–Cold War Defense Contraction," *National Tax Journal*, 2021, *74* (1), 9 – 43.

Kousser, Thad, Mathew Mccubbins, and Kaj Rozga, "When Does the Ballot Box Limit the Budget? Politics and Spending Limits in California, Colorado, Utah and Washington," in Elizabeth A. Graddy Elizabeth Garret and Howell E. Jackson, eds., *Fiscal Challenges: An Interdisciplinary Approach to Budget Policy*, Cambridge University Press, 2008.

Krumholz, Sam, "The Effect of District Attorneys on Local Criminal Justice Outcomes," *SSRN Electronic Journal*, 2020.

Lee, David S., Enrico Moretti, and Matthew J. Butler, "Do Voters Affect or Elect Policies? Evidence from the U. S. House," *The Quarterly Journal of Economics*, 2004, *119* (3), 807–859.

Liao, Yanjun and Carolyn Kousky, "The Fiscal Impacts of Wildfires on California Municipalities," *Journal of the Association of Environmental and Resource Economists*, 2022, *9* (3), 455 – 493.

Macartney, Hugh and John Singleton, "School Boards and Student Segregation," *Journal of Public Economics*, 2018, *164*, 165–182.

MacKay, Robert C., "Implicit Debt Capitalization in Local Housing Prices: An Example of Unfunded Pension Liabilities," *National Tax Journal*, 2014, *67* (1), 77–112.

McCarthy, Justin, "Perceptions of Increased U.S. Crime at Highest Since 1993," 2021.

McFall, Brooke Helppie, "Crash and Wait? The Impact of the Great Recession on the Retirement Plans of Older Americans," *American Economic Review*, 2011, *101* (3), 40–44.

Mello, Steven, "More COPS, less crime," *Journal of Public Economics*, 2019, *172* (C), 174–200.

Mueller-Smith, Michael, "The Criminal and Labor Market Impacts of Incarceration," 2015, *Working Paper.*

_ **and Kevin T. Schnepel**, "Diversion in the Criminal Justice System," *The Review of Economic Studies*, 2020, *88* (2), 883–936.

Novy-Marx, Robert and Joshua D. Rauh, "The Liabilities and Risks of State-Sponsored Pension Plans," *Journal of Economic Perspectives*, 2009, *23* (4), 191–210.

_ **and** _ , "Linking benefits to investment performance in US public pension systems," *Journal of Public Economics*, 2014, *116* (C), 47–61.

Ondrich, Jan and Alexander Falevich, "The Great Recession, Housing Wealth, and the Retirement Decisions of Older Workers," *Public Finance Review*, 2016, *44* (1), 109–131.

Pew, The Pew Charitable Trusts, "State Strategies to Detect Local Fiscal Distress," Technical Report 2016.

Piehl, Anne and Shawn Bushway, "Measuring and Explaining Charge Bargaining," *Journal of Quantitative Criminology*, 2007, *23*, 105–125.

Shem-Tov, Yotam, "Make or Buy? The Provision of Indigent Defense Services in the United States," *The Review of Economics and Statistics*, 2022, *1* (9).

Shi, Ying and John Singleton, "School Boards and Education Production," *American Economic Journal: Economic Policy*, 2022.

Shoag, Daniel, "The Impact of Government Spending Shocks: Evidence on the Multiplier from State Pension Plan Returns," *working paper*, 2011.

_ , "Using State Pension Shocks to Estimate Fiscal Multipliers since the Great Recession," *American Economic Review*, 2013, *103* (3), 121–24.

_ **, Cody Tuttle, and Stan Veuger**, "Rules Versus Home Rule—Local Government Responses to Negative Revenue Shocks," *National Tax Journal*, 2019, *72* (3), 543–574.

Sloan, CarlyWill, "How Much Does Your Prosecutor Matter? An Estimate of Prosecutorial Discretion," 2020, *Working Paper.*

Steven, Jonathan Schroeder David Van Riper Tracy Kugler Manson and Steven Ruggles, "IPUMS National Historical Geographic Information System: Version 17.0 [dataset]," `http://doi.org/10.18128/D050.V17.0` 2022. Minneapolis, MN: IPUMS.

Stith, Kate, "The Arc of the Pendulum: Judges, Prosecutors, and the Exercise of Discretion," *Yale Law Journal*, 2008, *117*, 1420–1497.

Thompson, Paul N., "Effects of fiscal stress labels on municipal government finances, housing prices, and the quality of public services: Evidence from Ohio," *Regional Science and Urban Economics*, 2017, *64*, 98–116.

Trounstine, Jessica, "Turnout and Incumbency in Local Elections," *Urban Affairs Review*, 2013, *49* (2), 167–189.

Tuttle, Cody, "Racial Discrimination in Federal Sentencing: Evidence from Drug Mandatory Minimums," 2019, *Working Paper.*

White, Ariel, "Misdemeanor Disenfranchisement? The Demobilizing Effects of Brief Jail Spells on Potential Voters," *American Political Science Review*, 2019, *113*, 311 – 324.

Yang, Lang (Kate), "Fiscal transparency or fiscal illusion? Housing and credit market responses to fiscal monitoring," *International Tax and Public Finance*, 2022, *29*, 1–29.

Yokley, Eli, "Most Voters See Violent Crime as a Major and Increasing Problem. But They're Split on Its Causes and How to Fix It," 2021.