Appendix: Supporting Information for *Political Power of Bureaucratic Agents*

Table of Contents

A Evidence Beyond NYC and Policing

A.1 Justification for Case Selection

To test my theory, the empirical case needs to justify the following theoretical and empirical scope conditions:

1. For theoretical purposes
   (a) A policy that police dislikes, that was implemented and that affects public service provision (otherwise, no plausible deniability for police)
   (b) Sufficient knowledge about policy among police
   (c) Police force with sufficient political autonomy and organization
2. For identification purposes
   (a) A contentious issue with conflict across politicians (within jurisdiction variation for identification purposes)
   (b) Fine-grained data that measures police effort and is arguably not a function of underlying time-varying confounders, e.g., crime
   (c) District elections, no at-large representation (otherwise, no within-jurisdiction variation)

Importantly, the FY2021 NYPD budget cut satisfies these conditions. The policy represented an unprecedented $1 billion reduction to the nation's largest police force, which police unions vocally opposed. The budget vote was highly contentious, with an unusual 17 council members voting against it compared to near-unanimous approval in prior years. New York City Council members represent geographically defined districts, which provides spatial variation in political alignment. Additionally, the NYPD is arguably the most professionalized police force in the US with significant political autonomy and strong union organization. Finally, the 911 response time data offers a direct measure of officer effort where police exercise considerable discretion, unlike downstream measures such as arrest rates that reflect both police and civilian behavior.

A.2 Minneapolis: Budget Cut and Police Shirking

To evaluate the broader applicability of my theory, I identified cities with similar or larger budget cuts compared to the NYC cut or contract negotiations affecting police in 2020, and evaluated each case against these scope conditions in Table A1. While the limited overlap between cases and conditions might suggest a narrow applicability of the leverage shirking theory, it is important to note that the exclusion of cases stems primarily from identification challenges rather than theoretical limitations. That is, the mechanism may still apply in these cities, it is just challenging to identify strategic shirking empirically.

Table A1: Evaluation of Additional Cases

Case	Reason	Theoretical suitability	Empirical suitability
Austin (Budget: -33% in 2020)	Unanimous decision in city council	✓	✗
Seattle (Budget: -11% in 2020)	little disagreement (8-1 decision); dissenting member wanted *larger* cut	✓	✗
Denver (Budget: -10% in 2020)	no disagreement on budget cut; 8-5 rejection of police contract in Sept. 2020, but solved by arbitrator in Oct. 2020	✓	✗
New Orleans (Budget: -8% in 2020)	both budget cut and union contract passed unanimously	✓	✗
Minneapolis (Budget: -8% in 2020)	7-6 vote on staffing & funding in Dec. 2020; micro data on police stops and response times granted through FOIA request	✓	✓
Las Vegas (Budget: -7% in 2020)	unanimous vote	✓	✗
Columbus (Budget: -6% in 2020)	at large elections	✓	✗
San Francisco (Budget: -5% in 2020)	10-1 vote, but dissent due to non-police issue	✓	✗
DC (Budget: -5% in 2020)	unanimous vote	✓	✗

I replicate my main analysis using the police budget in December 2020 in Minneapolis. Following George Floyd's murder, the city council decided to slash \$8 million (7.5%) from the Minneapolis Police Department (MPD). There was considerable controversy in the council about plans to hire more officers in future years. The City Council had initially planned to drop the force's authorized size to 750 officers starting in 2022, but reversed course by a narrow 7-6 vote to maintain the staffing level at 888.

To measure police effort, I rely on the number of officer-initiated stops made by MPD by day-district-precinct and response times for all 911 calls in 2020-2021.[1] Using a DiD strategy, I compare police behavior on day d and precinct p between misaligned and aligned districts c, i.e., districts where council members supported both the cut to staffing and funding (i.e., misaligned) and those where council members voted against the cut in staffing levels (i.e., aligned):

$$\log(\text{stops} + 1)_{cdp} = \alpha_c + \beta \text{misaligned}_c \times \text{after vote}_d + \gamma_d + \delta_p + \varepsilon_{cdp} \tag{1}$$

$$\text{response time}_{icdp} = \alpha_c + \beta \text{misaligned}_c \times \text{after vote}_d + \gamma_d + \delta_p + \varepsilon_{icdp} \tag{2}$$

The limited number of council districts (6 treated and 7 control districts) make inference challenging. Simple power analyses illustrate this. Without accounting for the clustered data structure, the minimal detectable effects with 80% power are 3% for the number of stops (0.05 standard deviations) and 1.35 minutes (0.02 standard deviations). However, accurately accounting for clustering in the data and treatment assignment significantly reduces power and increases the minimal detectable DiD effects substantially: 40% for the number of stops (0.64 standard deviations) and 11.4 minutes (0.11 standard deviations) for response times on average. This means I may be unable to detect more subtle but still policy-relevant effects that fall below these thresholds.[2] Given these power concerns, inferences need to be treated with caution in this analysis. I focus on the sign and size of the estimated DiD estimates and present robust standard errors without clustering as well as wild cluster bootstrap p-values following Roodman et al. (2019).

Overall, the estimates support the theory of leverage shirking and bolster the findings of the NYC case. As Table A2 shows, MPD officers reduced the number of stops by an additional 4% in misaligned districts compared to districts aligned with their preferences following the budget vote. The effect size is very similar when accounting for anticipation effects (i.e., excluding all days between George Floyd's murder and the budget vote on December 10, 2020) or excluding all arrests in the Powderhorn Park neighborhood where George Floyd died. Additionally, the effect seems largely driven by the intensive margin: Officers reduced the number of stops, conditional on making some stops in the district and precinct per day, by an additional 11% in treated districts relative to control districts. There is little evidence that they disproportionately reduced the probability of making *any* stop in the area. Conversely, the results in Table A3 suggest that response times disproportionately increased in misaligned districts after the budget vote by about 2 minutes. Again, the effect size is robust to excluding months right before the vote but after George Floyd's death and differentiating civilian-initiated from police-initiated calls.

[1] Stops data is taken from https://opendata.minneapolismn.gov/; response time data was obtained through an open records request (DR25_001925).

[2] I calculate the minimal detectable effect with 80% power as $(z_{1-alpha/2} + z_{1-beta}) \times \sqrt{V(\beta_{DiD})}$. I incorporate clustering in the variance of the DiD estimator: $V(\beta_{DiD}) = \sigma^2 \times \left(\frac{DE_{11}}{N_{11}} + \frac{DE_{10}}{N_{10}} + \frac{DE_{01}}{N_{01}} + \frac{DE_{00}}{N_{00}}\right)$, where $DE_{ij} = 1 + (m_{ij} - 1) \times ICC$ is the design effect adjustment using the intra-cluster correlation and average number of observations per cluster for each period-alignment-cell m_{ij} and N_{ij} are the number of observations in each cell.

Table A2: Treatment Effect Estimates for Number of Stops

	Base (1)	Restricted (2)	Excluding Powderhorn (3)	Intensive Margin (4)	Extensive Margin (5)
misalignment × after vote	-0.0406	-0.0511	-0.0395	-0.1093	-0.0002
	(0.0164)	(0.0264)	(0.0163)	(0.0217)	(0.0127)
Date FE	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓
Police Precinct FE	✓	✓	✓	✓	✓
Wild cluster bootstrap p-value	0.6550	0.7330	0.8840	0.5250	0.9940
Observations	20,468	15,008	20,468	7,350	20,468
Mean of DV	0.415	0.433	0.409	1.154	0.359
R^2	0.23489	0.24932	0.23755	0.42702	0.14373

Heteroskedasticity-robust standard-errors in parentheses.

Table A3: Treatment Effect Estimates for Response Times

	Base (1)	Severity (2)	Civilian initiated (3)	Restricted (4)
misalignment × after vote	1.789	1.959	2.072	1.982
	(0.4813)	(0.4797)	(0.6347)	(0.4978)
Date FE	✓	✓	✓	✓
District FE	✓	✓	✓	✓
Police Precinct FE	✓	✓	✓	✓
Call Importance FE		✓		
Wild cluster bootstrap p-value	0.1980	0.2260	0.1120	0.3200
Observations	746,102	746,102	554,321	553,608
Mean of DV	21.119	21.119	27.440	21.119
R^2	0.01348	0.02256	0.01212	0.01061

Heteroskedasticity-robust standard-errors in parentheses.

In addition to this quantitative evidence, there is ample qualitative evidence that the MPD targeted misaligned council members in the wake of the defund movement in Minneapolis. Following the protests, many residents began reporting delayed or nonexistent responses to emergency calls, prompting City Council members to publicly question whether police were engaging in deliberate slowdowns (Winter, 2020). Police Chief Medaria Arradondo attributed these delays to staffing shortages from officer departures and increased violent incidents that required longer officer engagement. However, some council members suspected their districts were specifically targeted because of their support for police defunding or restructuring initiatives. Council member Phillipe Cunningham, a fierce supporter of police reforms, reported that numerous constituents in his district received no police response to gunfire reports. When residents inquired about these delays, they were allegedly directed to contact Cunningham himself. Cunningham argued that police painted him as the villain: "It's my fault that they are not responding in a timely manner or at all" (Winter, 2020).

A.3 Other Bureaucracies

While a test of the theory of leverage shirking beyond police is beyond the scope of this paper and should be carefully evaluated in future research, the theoretical scope conditions of the argument travel to other street-level bureaucracies where service quality is observable but difficult to attribute to the bureaucrats or political decisions. For example, municipal water utility workers could adjust repair schedules based on council members' support for utility privatization; bus drivers and transit workers might leverage their direct interaction with passengers to shape perceptions of service cuts or route changes; and planning department staff could delay permitting to resist budget policies. Several case studies demonstrate how bureaucrats beyond law enforcement have strategically manipulated service provision to exert electoral pressure on elected officials.

A.3.1 Fire fighters in Washington, DC

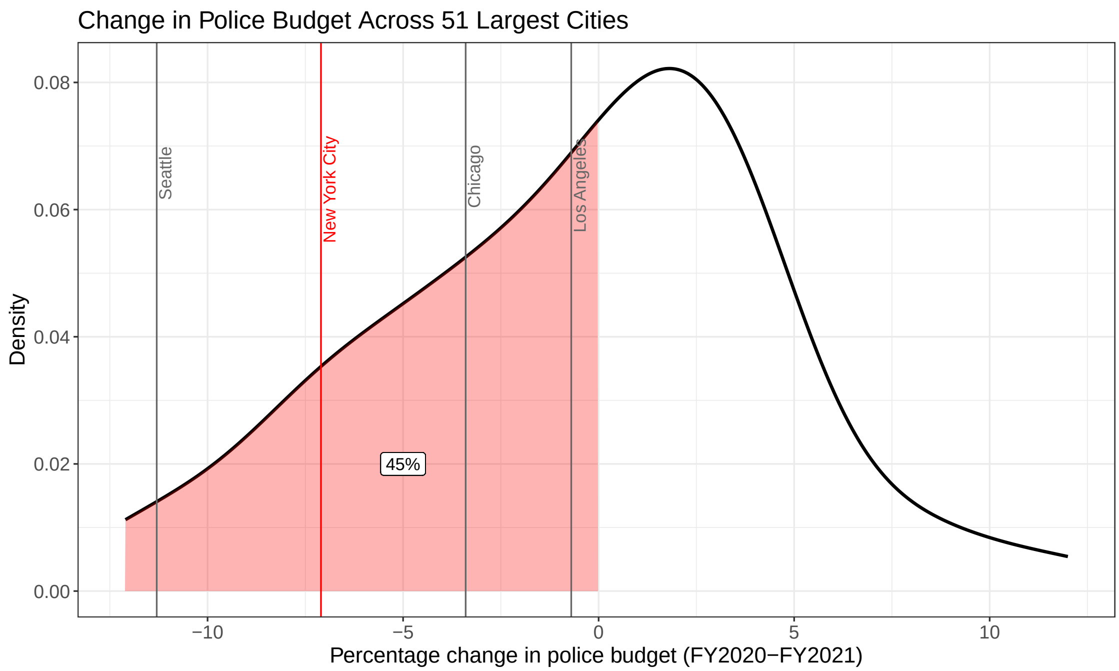
In November 2011, Fire Chief Kenneth Ellerbe proposed replacing the traditional scheduling system of 24-hour shifts followed by three days off with a "3-3-3" model (three 12-hour day shifts, three 12-hour night shifts, and three days off). The union, Local 36 of the International Association of Fire Fighters (IAFF), opposed this plan, citing concerns about increased fatigue and potential negative impacts on family life. As tensions escalated between the union and Fire Chief over policy differences, firefighters organized coordinated sick-outs in 2013. While typical weekly sick calls averaged between 20-30 firefighters, on August 18 alone, 83 firefighters reported illness (Hermann, 2013). The absences caused the department to require mandatory overtime of 67 firefighters, requiring them to work for 36 consecutive hours. DCFEMS officials called the illnesses suspicious, while the union said the illnesses showed the department had too few firefighters to cover the schedule due to Ellerbe's policies (NBC, 2013).

A.3.2 Sanitation workers in Staten Island

In 2010, sanitation workers in NYC repeatedly clashed with Mayor Bloomberg over budget cuts and cost-saving measures. Following a snow blizzard in December 2010, sanitation workers dragged their feet in snow removal efforts around the city to create a policy failure for Bloomberg. In Staten Island, for example, residents complained to representatives about abandoned or stuck plows and salt spreaders (Staten Island Live, 2010). In conversations with City Councilman Dan Halloran, sanitation workers revealed that "they were told [by supervisors] to take off routes [and] not do the plowing of some of the major arteries in a timely manner. They were told to make the mayor pay for the layoffs, the reductions in rank for the supervisors, shrinking the rolls of the rank-and-file." (Goldenberg, 2010). While the Bloomberg administration blamed residents for shoveling snow into streets that had already been plowed and for tying up 911 with non-emergency calls, the mayor soon became the public face of the failed handling of the storm.

B Background on Case: NYPD 2021 Budget Cut

Figure A1: Distribution of Police Budget Cuts Across Major US Cities in 2020



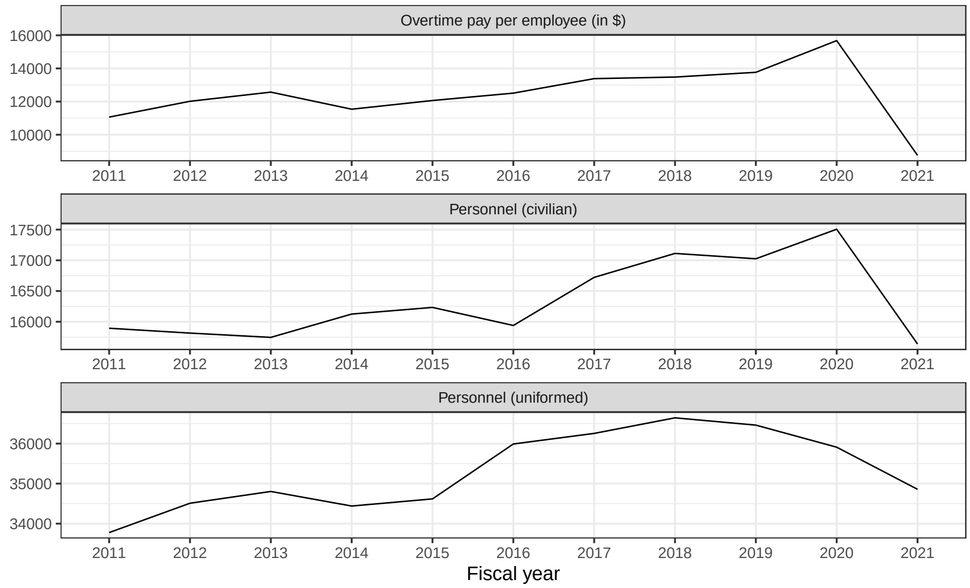
Note: The figure depicts changes in police budgets across all US state's largest cities, between fiscal years 2020 and 2021 (in percentages). Source: https://www.smartcitiesdive.com/news/calls-to-defund-the-police-are-upending-fy21-budgets-heres-how/581163/

Table A4: Summary Statistics - Covariates by Voting Behavior

	Vote on Budget Cut yes mean	Vote on Budget Cut no mean	difference est.	difference t-value
Council member characteristics				
Black candidate	37.50	23.53	-13.97	(-1.02)
Vote share last election	82.86	78.69	-4.18	(-0.89)
Win margin, last election	68.90	60.73	-8.17	(-0.92)
Term limited	59.38	64.71	5.33	(0.36)
Experience (in years)	6.09	5.59	-0.51	(-0.56)
Geographic characteristics (pretreatment)				
Vote share Biden 2020 [a]	79.81	67.74	-12.07*	(-1.95)
Share of white population [b]	26.47	46.71	20.25**	(2.57)
Share of black population [b]	27.95	14.17	-13.78*	(-1.95)
Share of hispanic population [b]	29.49	24.78	-4.71	(-0.82)
Share of female population [b]	52.84	52.30	-0.54	(-0.91)
Share of population over 65 [b]	12.16	12.53	0.37	(0.43)
Share of population over 18 [b]	78.28	78.60	0.33	(0.20)
Share of renter occupied households [b]	70.20	64.71	-5.48	(-1.05)
Number of George Floyd protests [c]	4.41	3.12	-1.29	(-0.97)
Number of violation complaints [d]	677.28	540.59	-136.69*	(-1.90)
Number of misdemeanor complaints [d]	2227.75	1621.88	-605.87***	(-2.88)
Number of felony complaints [d]	1330.91	1008.88	-322.02**	(-2.23)
Number of shootings [e]	15.81	9.29	-6.52	(-1.54)
Number of districts	32	17	49	

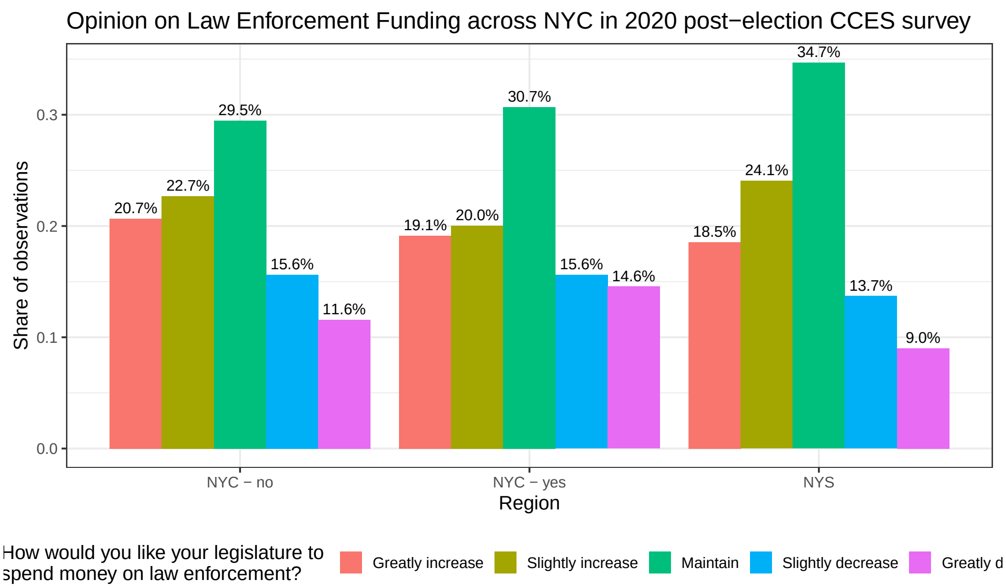
Sources: [a] Official Electoral Results, [b] Census Demographics, [c] Crowd Counting Consortium, [d] NYPD Complaint Data, [e] NYPD Shooting Incident Data.

Figure A2: Development of Personnel at NYPD



Note: The figure depicts NYPD resources from the FY2015, FY2020 and FY2021 Mayor's Management Reports (MMR), including paid overtime per employee, civilian personnel and uniformed personnel.

Figure A3: Preferences on Police Funding in 2020



Note: The figure depicts the distribution of survey responses regarding law enforcement funding from the 2020 post-election CCES survey, by type of NYC council districts and compared to citizens of New York State outside NYC. Missings are removed.

C Additional Results and Robustness

Figure A4: Weekly Average Differences between Districts by Agency

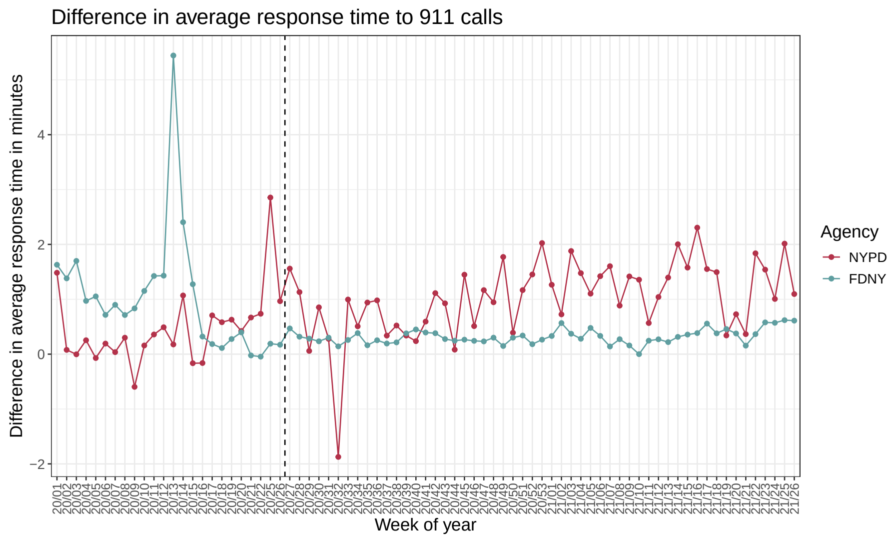
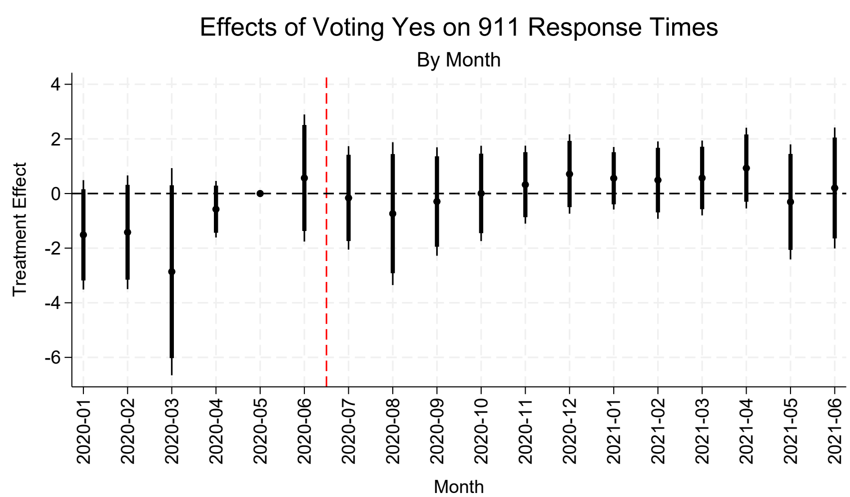


Figure A5: Monthly Treatment Effects



Note: Depicted are month-specific treatment effects, based on estimations of Equation (2) with 90% and 95% confidence intervals.

Figure A6: Call Length by Call Type

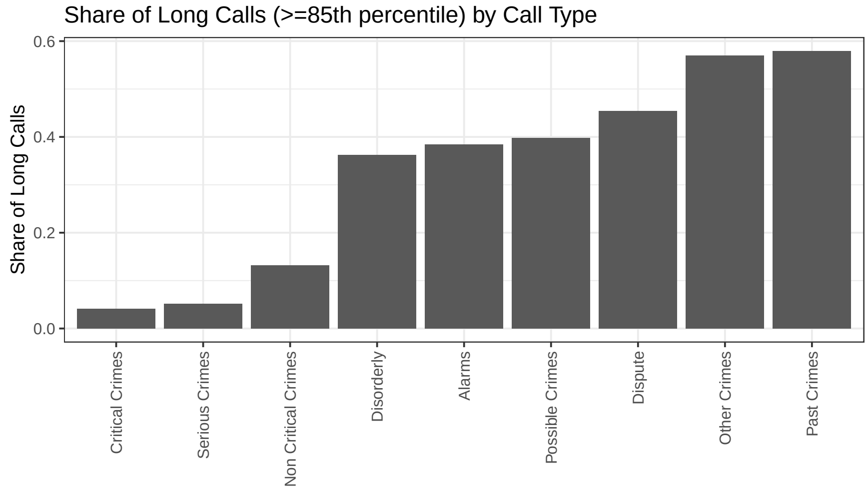


Table A5: Effect of Approving 2021 Budget on 911 Response Times, Including May 30 - June 15

	(1)	(2)	(3)	(4)
yesvote × postvote × NYPD	1.187*	1.208*	1.152*	1.218*
	(0.623)	(0.633)	(0.626)	(0.646)
NYPD	5.299***	5.749***	10.504***	-0.582
	(0.731)	(0.698)	(2.722)	(0.912)
yesvote × NYPD	0.170	-0.024	0.044	0.014
	(1.164)	(1.140)	(1.091)	(1.142)
postvote × NYPD	1.799***	1.785***	1.825***	1.839***
	(0.407)	(0.414)	(0.407)	(0.424)
yesvote × postvote	-0.667	-0.677	-0.707	-0.674
	(0.518)	(0.519)	(0.523)	(0.525)
District FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Police Precinct FE		✓	✓	✓
Daily call volume (log) × Agency			✓	
Call Importance FE × Agency				✓
Observations	9853758	9853736	9853736	9853736
Mean of DV	13.169	13.169	13.169	13.169
Adj. R^2	0.025	0.032	0.033	0.034

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Response time in minutes. Coefficients for yes vote$_c$ and after vote$_d$ absorbed by district and day fixed effects, respectively. Call importance fixed effects account for the two main levels of call importance for NYPD and FDNY calls: (1) Critical and serious crime incidents, life-threatening medical emergencies, and serious fires, (2) Non-critical crimes, non-crime incidents, non-life threatening medical emergencies, and low priority fire incidents. Cluster robust standard errors in parentheses, by district (49).

Table A6: Effect of Approving 2021 Budget on 911 Response Times, Simple DiD models

	(1)	(2)	(3)	(4)	(5)	(6)
		NYPD			FDNY	
yes vote × after vote	0.683*	0.603	0.699*	-0.811	-0.807	-0.806
	(0.393)	(0.389)	(0.398)	(0.520)	(0.520)	(0.525)
daily call volume (log)		-1.748***			0.225***	
		(0.281)			(0.072)	
District FE	✓	✓	✓	✓	✓	✓
Police Precinct FE	✓	✓	✓	✓	✓	✓
Date FE	✓	✓	✓	✓	✓	✓
Call Importance FE			✓			✓
Observations	7369246	7369246	7369246	2220981	2220981	2220981
Mean of DV	14.508	14.508	14.508	8.409	8.409	8.409
Adj. R^2	0.034	0.034	0.036	0.099	0.099	0.106

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Response time in minutes. Coefficients for yes vote_c and after vote_d absorbed by district and day fixed effects, respectively. Call importance fixed effects account for the two main levels of call importance for NYPD and FDNY calls: (1) Critical and serious crime incidents, life-threatening medical emergencies, and serious fires, (2) Non-critical crimes, non-crime incidents, non-life threatening medical emergencies, and low priority fire incidents.
Cluster robust standard errors in parentheses, by district (49).

Table A7: Effect of Approving 2021 Budget on 911 Response Times Excluding March 2020

	(1)	(2)	(3)	(4)
yes vote × after vote × NYPD	0.993**	0.998**	0.930**	1.010**
	(0.452)	(0.466)	(0.458)	(0.474)
NYPD	5.922***	6.339***	10.080***	-0.396
	(0.814)	(0.800)	(2.841)	(0.984)
yes vote × NYPD	0.349	0.171	0.228	0.203
	(1.313)	(1.301)	(1.249)	(1.305)
after vote × NYPD	1.178***	1.195***	1.170***	1.254***
	(0.300)	(0.316)	(0.307)	(0.320)
yes vote × after vote	-0.339	-0.346	-0.375	-0.344
	(0.249)	(0.250)	(0.257)	(0.253)
District FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Police Precinct FE		✓	✓	✓
Daily call volume (log) × Agency			✓	
Call Importance FE × Agency				✓
Observations	9007257	9007240	9007240	9007240
Mean of DV	12.968	12.968	12.968	12.968
Adj. R^2	0.025	0.032	0.033	0.034

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Response time in minutes. Coefficients for yes vote_c and after vote_d absorbed by district and day fixed effects, respectively. Call importance fixed effects account for the two main levels of call importance for NYPD and FDNY calls: (1) Critical and serious crime incidents, life-threatening medical emergencies, and serious fires, (2) Non-critical crimes, non-crime incidents, non-life threatening medical emergencies, and low priority fire incidents. Cluster robust standard errors in parentheses, by district (49).

Table A8: Effect of Approving 2021 Budget on 911 Response Times
Winsorized Response Times

	(1) 1-99 pct.	(2) 1-99 pct., by day
yes vote × after vote × NYPD	1.145**	1.203**
	(0.544)	(0.585)
NYPD	4.595***	4.582***
	(0.629)	(0.633)
yes vote × NYPD	-0.413	-0.481
	(0.971)	(0.961)
after vote × NYPD	2.067***	2.166***
	(0.368)	(0.396)
yes vote × after vote	-0.642	-0.685
	(0.473)	(0.513)
District FE	✓	✓
Police Precinct FE	✓	✓
Date FE	✓	✓
Observations	9590227	9590227
Mean of DV	12.294	12.339
Adj. R^2	0.041	0.042

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Response time in minutes. Coefficients for yes vote_c and after vote_d absorbed by district and day fixed effects, respectively. Cluster robust standard errors in parentheses, by district (49).

Table A9: Difference in Number of Calls by 2021 Budget Vote and Time

	Simple DiD (1)	Triple DiD (2)
yes vote × after vote × NYPD		-0.016
		(0.029)
NYPD		1.054***
		(0.090)
yes vote × NYPD		-0.004
		(0.119)
after vote × NYPD		0.011
		(0.020)
yes vote × after vote	-0.024	-0.016
	(0.025)	(0.013)
District FE	✓	✓
Police Precinct FE	✓	✓
Date FE	✓	✓
Observations	113700	212626
Mean of DV	3.715	3.263
Mean of untransformed DV	80.565	53.527
Adj. R^2	0.268	0.342

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Log number of calls by date, precinct and council district. Column (1) only includes NYPD calls. Cluster robust standard errors in parentheses, by district (49).

Table A10: Difference in Crime Incidents by 2021 Budget Vote and Time

	Crime calls	Serious crime calls	Shootings	Complaints
yesvote × postvote	-0.014	0.011	0.004*	-0.011
	(0.013)	(0.012)	(0.002)	(0.015)
District FE	✓	✓	✓	✓
Police Precinct FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Observations	114661	114661	114661	114661
Mean of DV	1.382	0.707	0.006	1.342
Adj. R^2	0.261	0.194	0.019	0.230

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Log number of calls for crimes in progress, log number of calls for serious crimes, log number of shootings, and log number of valid felony, misdemeanor, and violation complaints by date, precinct and council district. Cluster robust standard errors in parentheses, by district (49).

Table A11: Call Distance to NYPD Precinct Headquarters

	(1)	(2)
yes vote	-334.571*	-323.317*
	(185.173)	(184.251)
yes vote × after vote		-16.546
		(13.410)
Date FE	✓	✓
Police Precinct FE	✓	✓
Observations	9160217	9160217
Mean of DV	1265.811	1265.811
Adj. R^2	0.402	0.402

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Cluster robust standard errors in parentheses, by district (49).

Table A12: Difference in Crime Clearance Rates by 2021 Budget Vote and Time

	All	Felony	Misdemeanor	Violation
	(1)	(2)	(3)	(4)
yes vote × after vote	-0.0056	-0.0109	-0.0100	-0.0001
	(0.0106)	(0.0078)	(0.0120)	(0.0068)
District FE	✓	✓	✓	✓
Police Precinct	✓	✓	✓	✓
Date	✓	✓	✓	✓
Observations	135,656	135,656	135,656	135,656
Mean of DV	-0.588	-0.276	-0.434	-0.370
R^2	0.15941	0.09391	0.11533	0.13933

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Difference in log number of arrests and log number of crime complaints, i.e., log clearance rate. Level of observation: Date-district-precinct. Cluster robust standard errors in parentheses, by district (49).

Table A13: Effect of Approving 2021 Budget on 911 Response Times Accounting for Protests

	(1) Simple DiD	(2)	(3) Triple DiD	(4)
yes vote × after vote × NYPD		1.432**	1.403**	1.379**
		(0.680)	(0.673)	(0.671)
NYPD		5.257***	5.213***	5.372***
		(0.699)	(0.694)	(0.709)
yes vote × NYPD		-0.250	-0.255	-0.299
		(1.108)	(1.110)	(1.101)
after vote × NYPD		2.274***	2.404***	2.426***
		(0.439)	(0.446)	(0.446)
yes vote × after vote	0.684*	-0.755	-0.724	-0.707
	(0.393)	(0.563)	(0.553)	(0.552)
# of protests (log)	0.298	0.090	-5.511***	
	(0.761)	(0.445)	(1.187)	
after vote × # of protests (log)	-0.314		7.820***	
	(0.970)		(1.786)	
NYPD × # of protests (log)			7.558***	
			(1.919)	
after vote × NYPD × # of protests (log)			-10.448***	
			(2.875)	
# of protests (log) (June 2020)				0.490
				(0.920)
after vote × # of protests (log) (June 2020)				1.102**
				(0.541)
NYPD × # of protests (log) (June 2020)				-1.240
				(2.051)
after vote × NYPD × # of protests (log) (June 2020)				-1.628
				(1.043)
District FE	✓	✓	✓	✓
Police Precinct FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Observations	7369246	9590227	9590227	9590227
Mean of DV	14.508	13.095	13.095	13.095
Adj. R^2	0.034	0.032	0.032	0.032

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Response time in minutes. Coefficients for yes vote$_c$ and after vote$_d$ absorbed by district and day fixed effects, respectively. Cluster robust standard errors in parentheses, by district (49).

Table A14: Effect of Approving 2021 Budget on 911 Response Times
Robustness to Compound Treatments

	(1)	(2)	(3)	(4)
	Triple DiD		Simple DiD	
yes vote × after vote × NYPD	1.16*	1.19*		
	(0.67)	(0.68)		
yes vote × after vote	-0.51	-0.40	0.66	0.79**
	(0.52)	(0.55)	(0.40)	(0.39)
NYPD	6.06***	7.39***		
	(0.96)	(2.54)		
yes vote × NYPD	-0.44	0.04		
	(1.11)	(1.01)		
after vote × NYPD	3.28***	0.57		
	(0.67)	(1.18)		
white councilor × after vote × NYPD	-1.88***			
	(0.67)			
white councilor × after vote	1.64***		-0.30	
	(0.48)		(0.38)	
white councilor × NYPD	-1.53			
	(1.19)			
Biden vote share × after vote × NYPD		2.34		
		(1.66)		
Biden vote share × after vote		-3.59***		-1.12
		(1.09)		(1.26)
Biden vote share × NYPD		-2.92		
		(3.28)		
District FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Police Precinct FE	✓	✓	✓	✓
Observations	9590227	9590227	7369246	7369246
Mean of DV	13.095	13.095	14.508	14.508
Adj. R^2	0.032	0.032	0.034	0.034

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Response time in minutes. Coefficients for yes vote$_c$ and after vote$_d$ absorbed by district and day fixed effects, respectively. Cluster robust standard errors in parentheses, by district (49).

Table A15: Effect of Approving 2021 Budget on Probability of Officer-Initiated Calls

	(1)	(2)	(3)	(4)
	Response Time = 0		Response Time < 0.15	
yes vote × after vote × NYPD		0.013		0.002
		(0.014)		(0.009)
NYPD		-0.011***		0.553***
		(0.002)		(0.011)
yes vote × NYPD		-0.003		-0.007
		(0.002)		(0.017)
after vote × NYPD		0.354***		-0.006
		(0.008)		(0.007)
yes vote × after vote	0.011	-0.003***	0.001	-0.001
	(0.014)	(0.001)	(0.009)	(0.001)
District FE	✓	✓	✓	✓
Police Precinct FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Observations	7369246	9590227	7369246	9590227
Mean of DV	0.250	0.194	0.569	0.439
Adj. R^2	0.315	0.323	0.039	0.251

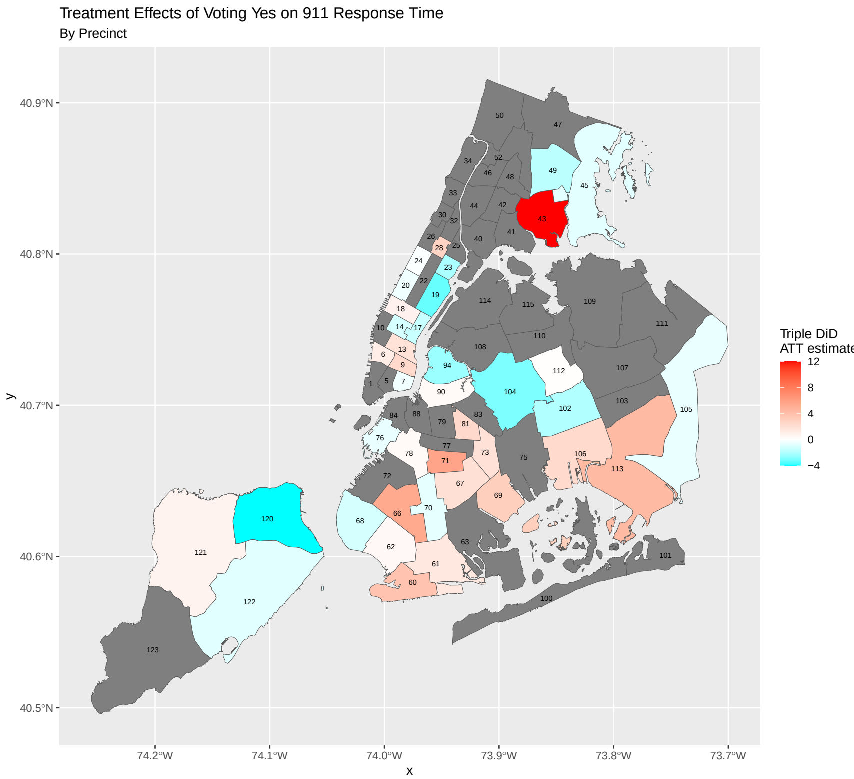
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Dummy for zero or < 0.15 response time. Columns (1) and (3) only include NYPD calls. Coefficients for yes vote_c and after vote_d absorbed by district and day fixed effects, respectively. Cluster robust standard errors in parentheses, by district (49).

Table A16: Effect of Approving 2021 Budget on 911 Response Times, Excluding Zero Response Time Calls

	Simple DiD	Triple DiD		
	(1)	(2)	(3)	(4)
yes vote × after vote × NYPD		2.396*	2.350*	2.468*
		(1.317)	(1.318)	(1.344)
NYPD		5.249***	7.332**	-7.494***
		(0.702)	(3.097)	(1.099)
yes vote × NYPD		-0.237	-0.213	-0.184
		(1.113)	(1.063)	(1.118)
after vote × NYPD		10.543***	10.549***	10.947***
		(0.834)	(0.838)	(0.859)
yes vote × after vote	1.548*	-0.823	-0.856	-0.821
	(0.895)	(0.569)	(0.574)	(0.576)
District FE	✓	✓	✓	✓
Police Precinct FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Daily call volume (log) × Agency			✓	
Call Importance FE × Agency				✓
Observations	5523493	7727607	7727607	7727607
Mean of DV	19.355	16.252	16.252	16.252
Adj. R^2	0.052	0.056	0.057	0.061

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Response time in minutes. Coefficients for yes vote_c and after vote_d absorbed by district and day fixed effects, respectively. Cluster robust standard errors in parentheses, by district (49).

Figure A7: Average Marginal Effects of Yes Vote by Precincts



Note: Depicted are ATT estimates from regressions within each precinct separately. Consequently, only precincts that straddle council districts with opposing budget votes are included.

Figure A8: Trends in Amount of 911 NYPD Calls across Districts

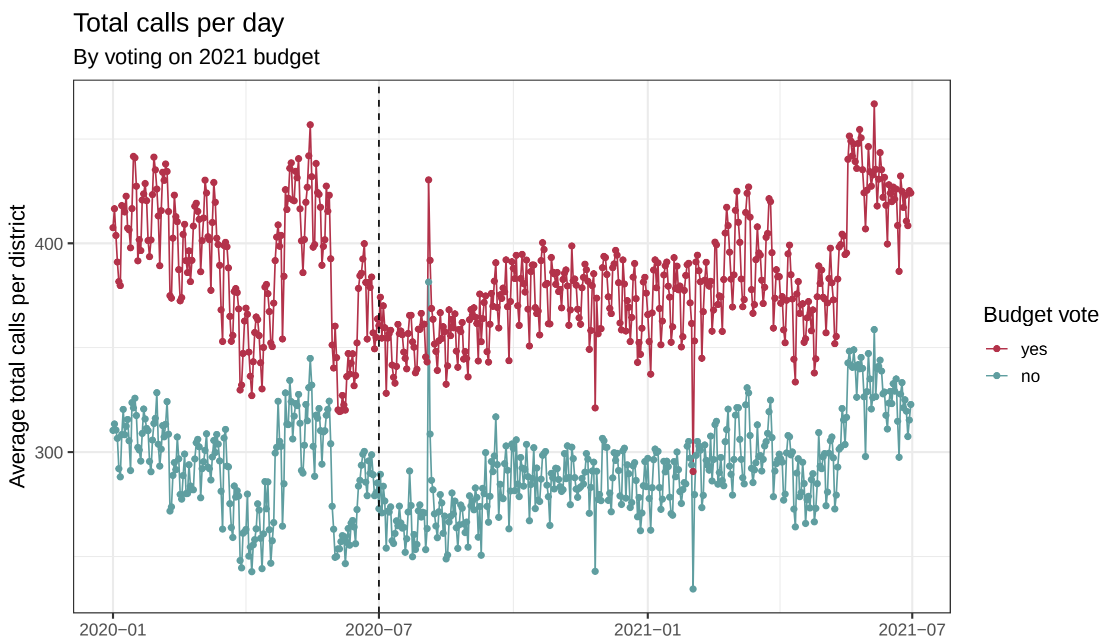


Figure A9: Distribution of 911 Call Types, by Period and District

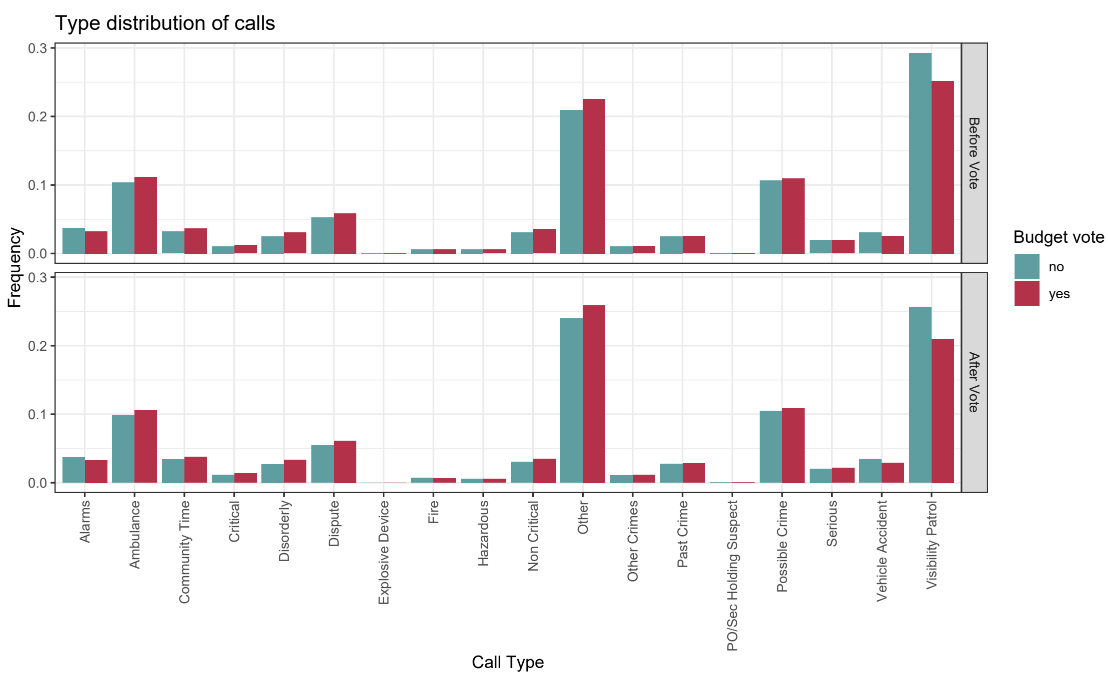
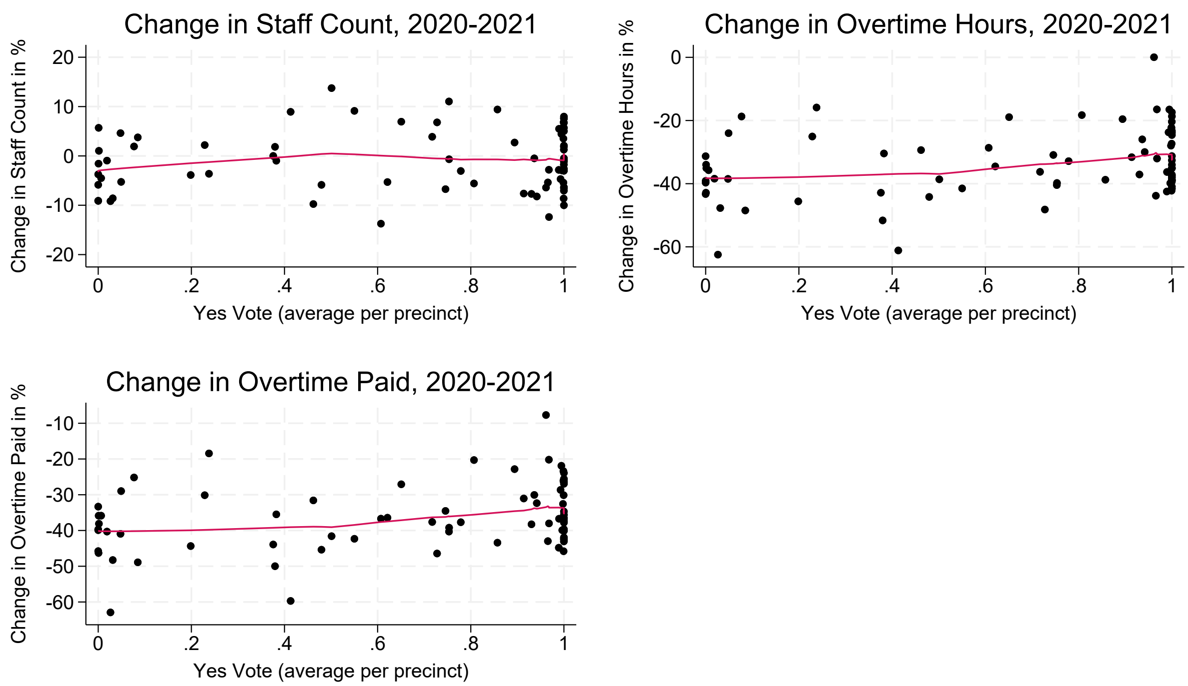


Figure A10: Correlation of Treatment and Change in Resources by Precinct



D Spatial Difference-in-Discontinuities Design

The DiD design crucially hinges on the comparability of treated and control districts in the entire city, across agencies and over time, i.e., the validity of the parallel trends assumption. As the main results indicate, this might be violated and complicated by the fact that police could shirk shortly *before* the vote in the hope to influence council members' voting behavior. To leverage more local variation in a-priori similar neighborhoods, I therefore supplement the analysis with a spatial difference-in-discontinuities design. I use a spatial RDD design to compare NYPD response times in close proximity to the council district borders that separated yes and no voting members (see Figure A11). For each 911 call, I calculate the minimum distance to a separating border to construct the running variable. To provide estimates for the changes in these RDD estimates before and after the vote, I split my sample along the date of the budget vote.[3] For both time periods, the resulting model is estimated as follows:

$$\begin{aligned}\text{response time}_{icpd} = \alpha + \tau \text{yes vote}_c + \beta_{-}\text{distance}_{icpd} + \beta_{+}\text{yes vote}_c \times \text{distance}_{icpd} \\ + \eta_p + \varepsilon_{icpd}\end{aligned} \quad (3)$$

where response time_{icpd} is the response time of call i in district c and day d, yes vote_c is an indicator equal to 1 if council member of district c voted in favor of the budget cut. distance_{icpd} represents the minimum distance of call i to the border distinguishing these two categories of districts, and contains only units $\text{distance}_{icpd} \in [-h; h]$, where $-h$ and h denote the MSE-optimal bandwidths to the left and right of the border, respectively. The model is estimated using local linear regression with a triangular kernel (Calonico et al., 2014). NYPD precinct fixed effects again account for systematic differences in response times across police management units. I use Monte Carlo simulations to provide confidence intervals of the difference in RDD estimates (King et al., 2000).

A few clarifying comments are warranted. Like all spatial RDD settings that rely on administrative borders, estimates of τ likely suffer from compound treatment problems, since many characteristics beyond a council member's vote change discontinuously along district borders, such as road quality or demographics. Yet, this is less problematic in a *difference*-in-discontinuities design. To the extent that these characteristics and their effect on NYPD response times stay constant across the periods before and after the vote, the *difference* in the RDD treatment effects remains unbiased. Yet, if other determinants of NYPD response times change over time along the separating border, the difference in RDD estimates represents an estimate for the heterogeneity in the treatment effect across periods, rather than a full-fledged causal moderation analysis. To alleviate these concerns, I estimate RDD estimates where I match observations across periods using coarsened exact matching on either side of the cutoff on relevant covariates, including call type and the number of calls per day on the zip code level. Figure A12 and Figure A13 show the resulting balance in these covariates after matching.

[3] As before I exclude dates affected by the George Floyd protest (May 29 - June 15, 2020). Additionally, to avoid concerns about anticipatory police behavior, I also exclude calls between June 16 - June 30, 2020.

Figure A11: RDD Sample

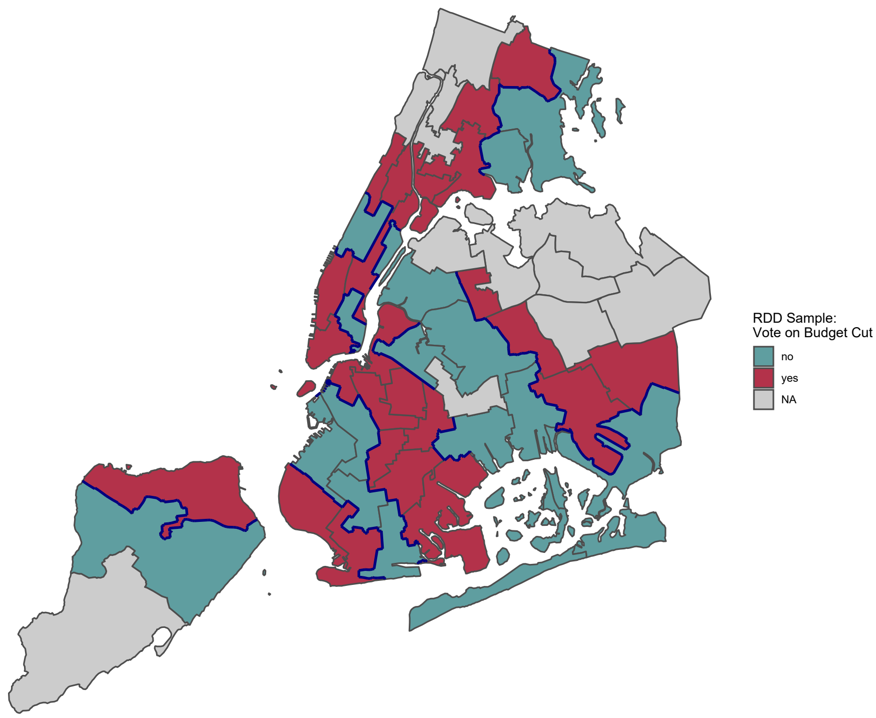


Figure A12: Balance of Matched RDD Sample - Major Call Types

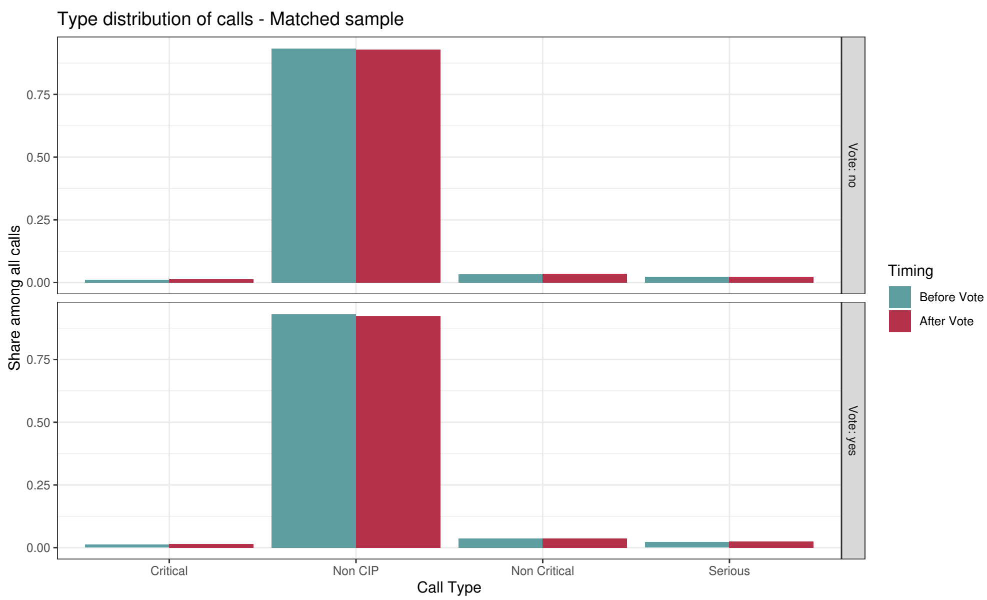


Figure A13: Balance of Matched RDD Sample - Daily Call Volume by Zip Code

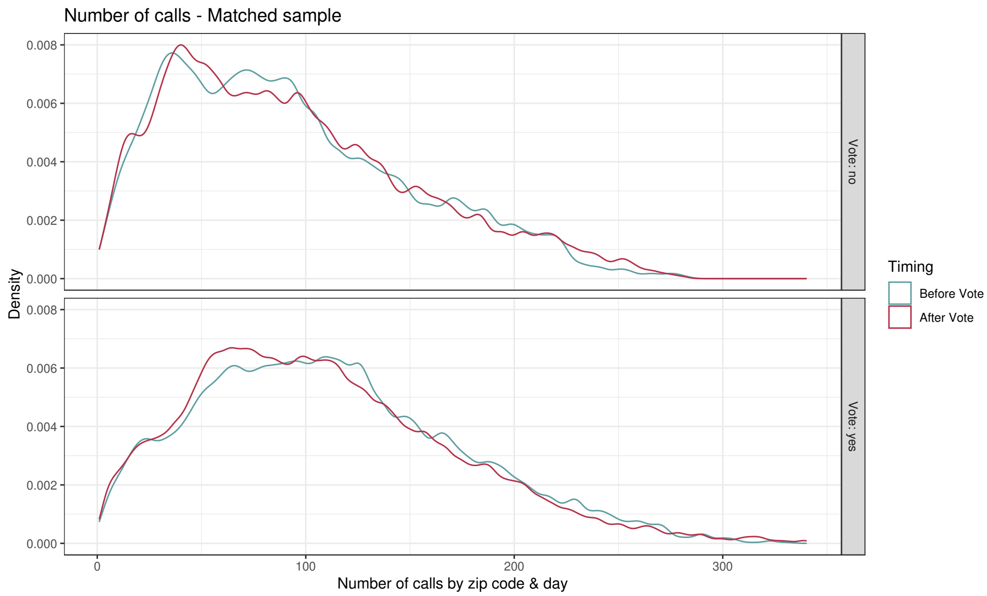


Table A17 shows the results. Interestingly, the negative RDD estimates in both periods suggest that NYPD officers respond faster to calls in treatment districts (yes votes) compared to neighboring control districts (no votes), both before and after the vote. This might be attributed to systematic differences in these neighborhoods that determine response times, including traffic, road quality etc.[4] More importantly for my argument, the difference in the RDD estimates is positive and significant. In line with previous results, the model suggests that for neighborhoods in close proximity to the district borders NYPD slowed down by about 1.5 minutes per call in yes voting districts relative to no voting districts after the budget vote.

Table A17: Effect of Approving 2021 Budget on 911 Response Times
Spatial Difference-in-Discontinuities

	Before Vote	After Vote	Difference
yes vote (robust bias-corrected)	-2.813	-1.359	1.453
	(-3.189; -2.436)	(-1.627; -1.092)	(0.993; 1.916)*
Precinct FE	✓	✓	
Matched Sample	✓	✓	
Kernel	Triangular	Triangular	
Bandwidth	mserd	mserd	
BW_est	231.767	161.985	
Obs_left	599,725	1,411,730	
Obs_right	1,254,137	2,844,357	

Dependent variable: Response time in minutes. 95% confidence intervals shown in parentheses. * 95% CIs from Monte Carlo simulations.

[4] Table A18 indicates that calls in yes-voting parts of the RDD sample are slightly closer to the precinct headquarter, thus presumably shortening the amount of travel necessary.

Table A18: Call Distance to NYPD Precinct Headquarters RDD Sample (within 200 meter bandwidth)

	(1)	(2)
yes vote	-72.57	-70.63
	(48.42)	(50.64)
yes vote × after vote		-2.72
		(12.89)
Police Precinct FE	✓	✓
Date FE	✓	✓
Observations	1,333,601	1,333,601
Mean of DV	1129.69	1129.69
Adj. R^2	0.78	0.78

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Cluster robust standard errors in parentheses, by district (49).

E Changes in Public Safety Concerns

In this section, I study how citizens' concerns about crime diverged across types of council districts after the budget cut. I use micro-level data from the monthly Gallup Social Series (2019-2023), which includes a question on what issue respondents perceive to be the most important problem facing the country today. Information about the zip code of respondents allows me to match respondents in New York City to council districts.[5] Any interpretation of the following results requires considerable caution since restricting the Gallup data to only observations in the relevant neighborhoods of New York City yields a small number of observations and these survey data are by no means representative on the level of the council district. I estimate a simple difference-in-differences model:

$$MIP(crime)_{ijt} = \alpha + \beta \text{yes vote}_j \times \text{post vote}_t + \delta_j + \gamma_t + \mathbf{X'_{ijt}}\rho + \varepsilon_{ijt} \quad (4)$$

where $MIP(crime)_{it}$ is a dummy for whether respondent i in district j and month t mentions that crime is one of the top three most important issues in the country at the time. δ_j and γ_t are council and month fixed effects, respectively. $\mathbf{X_{ijt}}$ are respondent-level controls for partisanship and race.

Table A19 presents the results of the difference-in-differences design, and Figure A14 depicts average predicted probabilities based on column (2) of Table A19. The results suggest that citizens in NYC were disproportionately more concerned about crime after the budget cut in yes-voting than in no-voting districts. While these patterns are only descriptive and may be driven by a more general shift in the political environment, they are in line with the idea that police may play with the citizens perceptions of public safety as a result of their day-to-day service provision.

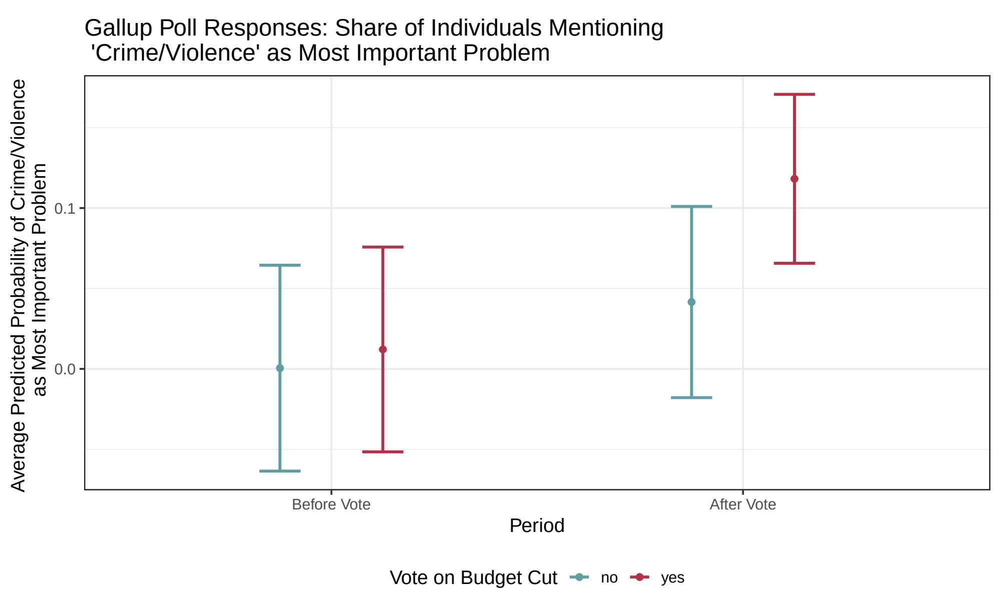
[5]Since zip codes are not perfectly subsumed in council districts, I match each zip code to the council district that accounts for the majority of its geographic area.

Table A19: Effect of Approving 2021 Budget on Crime Concerns

	(1)	(2)	(3)
yesvote × postvote	0.06*	0.06*	0.07*
	(0.03)	(0.03)	(0.03)
postvote	0.02		
	(0.03)		
yesvote	−0.03		
	(0.02)		
Council districts FE		✓	✓
Month FE		✓	✓
Individual controls			✓
Num. obs.	808	808	808
N Clusters	49	49	49

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Dependent variable: Dummy for indicating 'Crime/Violence' as MIP. Standard errors clustered by council districts. Individual controls include partisanship and race.

Figure A14: Predicted Probabilities of Indicating Crime as 'Most Important Problem'



F Impact on Candidate Vote Share

In this section, I provide some correlational evidence suggesting that council members opposed to police interests incurred electoral costs in the 2021 municipal elections relative to aligned council members. For this exercise I collect administrative data on election results on the election district level (i.e. the smallest electoral unit within a council district) for the 2017 and 2021 city council elections from the NYC Board of Elections.[6] For each electoral district and election I then calculate the vote share for council members voting on the 2021 budget.

Several aspects complicate this analysis. First, since I am interested in whether incumbents lost votes due to their votes on the 2021 budget, my sample is restricted to council members who ran in both elections and to districts where general/primary elections took place in both years. Another caveat arises due to a change in NYC's electoral system in 2021. New York City switched to rank-choice voting (RCV) for primary elections, allowing voters to rank up to five candidates for each race. Earlier elections were conducted under a standard first-past-the-post format. This implies a slight modification of my outcome variable, since vote shares are no longer simple to estimate. To calculate an incumbent's vote share that is comparable to my measure for the 2017 elections, I use individual-level cast vote records to compute the share of voters within a precinct who ranks each candidate as their top choice. This measure is easy to grasp and relatively analogous to vote shares in a first-past-the-post system.

I then estimate the following first-difference model:

$$\Delta voteshare_{ie} = \alpha + \beta \text{misalignment}_i + \varepsilon_{ie} \quad (5)$$

where I regress a council member i's difference in their vote share in electoral district e between 2017 and 2021 on whether they voted yes as opposed to no on the 2021 budget. Since there is a very small number of council districts in this model (9 for the primary elections, 11 for the general elections), I use heteroskedasticity-robust standard errors in parentheses and also present wild cluster bootstrap p-value following Roodman et al. (2019).

Table A20: Effect of Approving 2021 Budget on 2021 Election Vote Shares

	Primary	General
misalignment	−0.16	−0.09
	(0.01)	(0.01)
Mean of DV	0.055	0.128
Wild cluster bootstrap p-value	0.14	0.57
Num. obs.	869	1059
Adj. R^2	0.09	0.03

Dependent variable: Δ in vote share for incumbent on electoral district level. Heteroskedasticity-robust standard errors in parentheses. Bootstrap p-value refers to the coefficient on misalignment and is computed using the cluster wild bootstrap procedure of Roodman et al. (2019).

The results in Table A20 suggest that approving the 2021 budget cut was associated with a

[6] https://vote.nyc/page/election-results-summary

reduction in council member's vote shares. In the Democratic primary elections, where most of the electoral competition takes place in NYC, incumbents who supported the budget cut lost 16 percentage points more than council members opposing the substantial cut. In fact, two of the seven council members in favor of the budget cut in this sample lost their primary elections all together – a rare event for incumbents in NYC's Democratic primaries. Given the important caveats of this analysis and the imprecision, these estimates do not allow for causal inferences and conclusive interpretations. Yet, they provide some correlational evidence that council members who acted contrary to police interests during the 2021 budget vote might have incurred some electoral costs in the upcoming city elections.

References

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs". *Econometrica* 82 (6): 2295–2326.

Goldenberg, Sally. 2010. "Sanitation Department's slow snow cleanup was a budget protest". New York Post, Dec. 30, 2010. https://nypost.com/2010/12/30/sanitation-departments-slow-snow-cleanup-was-a-budget-protest/.

Hermann, Peter. 2013. "Deep management troubles overshadow D.C. fire department efforts to roll out good news". Washington Post, Aug. 26, 2013. https://www.washingtonpost.com/local/deep-management-troubles-overshadow-dc-fire-department-efforts-to-roll-out-good-news/2013/08/26/f99ee292-0b48-11e3-8974-f97ab3b3c677_story.html.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analysis: Improving Interpretation and Presentation". *American Journal of Political Science* 44 (2): 347–361.

NBC. 2013. "Changes to D.C. Fire Sick Leave Following Death". NBC Washington, Jan. 16, 2013. https://www.nbcwashington.com/news/local/changes-to-dc-fire-sick-leave-following-death/2070032/.

Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb. 2019. "Fast and Wild: Bootstrap Inference in Stata Using boottest". *The Stata Journal* 19 (1): 4–60.

Staten Island Live. 2010. "Sanitation employees posing work slowdown? Residents say yes; city, union say no". Staten Island Advance, Dec. 28, 2010. https://www.silive.com/news/2010/12/sanitation_employees_posing_wo.html.

Winter, Deena. 2020. "Data: A slower, less proactive Minneapolis Police Department since George Floyd killing". Minnesota Reformer, Sept. 15, 2020. https://minnesotareformer.com/2020/09/15/data-show-a-slower-less-proactive-minneapolis-police-department-since-george-floyd-killing/.