Assessing the impact of police officer line-of-duty deaths and force size on civilian arrests

Srinidhi Narayanan

March 28, 2024

Abstract

The death of a police officer in the line of duty can be traumatic for the fallen officer's colleagues, and may elicit a response with downstream consequences — does the death of an officer in the line of duty affect how the force polices in the aftermath? This paper investigates this question by assessing the impact of an officer death on arrest numbers in the officer's municipality for up to nine months post-death with a difference-in-differences approach. Furthermore, the paper investigates the role of police force size in this impact with a triple difference estimator. I find that force size has no significant effect on arrests. I find that the death of an officer in the line of duty has a significant effect in the month of the death, and no significant effects in the nine months post-death. I find that the line of duty death also plays a significant effect as a term in the triple difference regression. This suggests that police force size may be correlated with the death of an officer in the line of duty in subtle ways not reflected in the dataset.

1 Introduction

Policing is a critical social and political issue globally. This is even more the case in the United States, a country which incarcerates disproportionate to its size and in which, compared to other democracies, civilian-officer relationships are unusually strained [1] [2]. The issue of policing intersects with broader themes of civil liberties, state authority, and public trust, and recent high-profile cases of police brutality have sparked nationwide calls for comprehensive police reform. It is therefore imperative that we develop a robust understanding of policing mechanisms and the drivers of police behavior.

The death of a police officer in the line of duty (LODD - Line Of Duty Death) is a shock that is oftentimes traumatic for the fallen officer's colleagues [8], and which may elicit a response in the form of changed policing behavior [4]. If an officer is killed in the line of duty, it is possible that other officers may engage in risk-mitigating behavior in the aftermath, being more reluctant to police or unwilling to interact with the public, causing arrest numbers to decrease. On the other hand, officers might become hyper-vigilant, over-policing and possibly driving arrest numbers upward. A body of work studying this question already exists — most notably [3] — but studies actually vary in their findings, even up to the sign of the impact.

Additionally, it is possible that the size of the affected police force is a contributor to this impact. In a smaller police force, the death of an officer may feel more personal to colleagues and may have an outsize impact on the change in arrest numbers; the size of the force may also indicate the level of support or resources the department experiences from the city, with a larger force benefitting from more support. Support and resources could reduce the impact of the officer death [5].

In summary, my paper will investigate the two-fold question "How does the death of an officer in the line of duty affect future policing behavior? And, does this effect vary based on police force size?" Though the former question has been answered before with the same data source I use in [3], I introduce a data sample restriction that I hypothesize will provide more accurate estimates of the coefficients derived in that paper. My paper also adds a (to my knowledge) previously-unexamined heterogeneity study with the investigation of the effect of police force size.

To this end, I examine FBI datasets with samples from 20,000 municipalities over a five year period from 2013-2017, and my empirical strategy uses a difference-in-differences design that exploits the staggered occurrence of line-of-duty deaths across municipalities and time. The initial analysis finds a significant effect in the exact month of the killing, and no significant treatment effects beyond month 0. I then use a triple-difference estimator [9] to examine the additional impact of force size, also finding no significant effect.

Section 2 of this paper highlights the datasets and data preparation. Section 3 describes the empirical framework and key assumptions of the difference-in-differences analysis; and summarizes and contextualizes key results. Section 4 examines possible future steps.

2 Data

2.1 Datasets

I use FBI datasets, further cleaned, concatenated, and available with more granularity via openICPSR [6] [7]. All data is available by municipality by month for the years 2013 to 2017. I use the Law Enforcement Officers Killed and Assualted (LEOKA) dataset, and for the arrest numbers, the Uniform Crime Reporting (UCR) dataset. I also make use of employment data that is contained in the LEOKA dataset.

Each dataset initially contains data on around 20,000 municipalities. LEOKA contains columns on the breakdowns of various types of assault; demographic information on the assailant(s) and victim(s); the populations of each municipality in each year; and as stated above, information on the numbers of male and female federal and civilian officers in each year. UCR contains columns pertaining to arrests of different types, weapons involved in those arrests, etc.

2.2 Sample Restriction

It is evident from the range of populations represented that municipalities vary significantly in size; it is reasonable to assume they would vary in many other ways as well. Adding place fixed effects to account for demographic and other differences is one option, but an additional sample restriction could act to further control the analysis. I choose to restrict the data to only municipalities where an officer was at least shot at some point in the five-year period, and consider the treatment to be an officer being killed. With this control, the officer's being killed becomes a quasi-random occurrence, and is thus a natural variation I make use of. After adding the sample restriction, the size of the dataset reduces by about 50 percent to around 10,000 municipalities.

An overview of quantities of interest after applying the sample restriction, is here:

Table 1: Summary Statistics

	Variables			
	Mean	S.D.	Min	Max
Population	33419	100625	0	4007905
Total Employee Officers	61.5	252.3	0	9988
Officers Killed	.0005	.03	0	5
Total Officers Assaulted	.62	4.05	0	566
Total Arrests	21.7	80.9	-55	20000

For all quanitites of interest represented above, the standard deviation is much larger than the mean. This is due to some very large municipalities' skewing the data — the presence of these large municipalities is evidenced by the "Max" column.

2.3 Preprocessing

After retrieving the datasets from the openICPSR repositories, I apply some preprocessing before performing the analysis. I add a total assaults variable and an assault indicator variable; and encode the months from January of the first year as a 1 to December of the last year as a 60.

Since I am interested in the time trajectory of officer behavior, I want to view indicators for several months post-treatment, where the treatment is an officer killing in a municipality in a month. I would also like to verify that there is no pre-treatment effect, so I wish to consider pre-treatment months as well. So, I create nine indicator variables to indicate "officer killed 1 month prior", "officer killed 2 months prior", ..., "officer killed nine months prior". When an officer is killed in a municipality j in month t, the dummy column for "1 month prior" is updated with a 1 in the row corresponding to municipality j, month $t+1$; and so forth. Likewise, I create 1, 2, and 3 month pre-treatment dummies. The assumption in this creation is that the effect of the death, if any exists, will be apparent in nine months, and that there is no treatment effect after nine months.

I transform the arrests (dependent variable) column to a log of the arrests, since the interpretation of the regressions will then be with respect to percent changes; using raw arrest numbers could yield problems where results are skewed by some municipalities with very large or very small populations.

Finally, since I am interested in police force size, I calculate sizes, measured two ways: the first is a notion of "absolute" force size, and is a 1/0 indicator for large/small police force based on whether the force is larger or smaller than the median force size across all municipalities. I also calculate a measure that represents force size relative to the municipality's population with the metric $\frac{\text{force employment}}{\text{population of municipality}}$.

The resultant data is a panel dataset with time variable being the numerically encoded month, and place variable being the municipality.

3 Empirical Strategy

I employ a difference in differences strategy for this analysis, making some key assumptions in the process.

The foremost assumption is that of parallel trends: in the absence of a police killing, the arrest trends in untreated and treated places are similar. An earlier paper [3] investigating the question of trends post police-killing conducts several robustness and balance tests that appear to verify this assumption. In the results described below, it also becomes clear that the pre-treatment coefficients do not exhibit trends, which, while not equivalent, is consistent with this assumption.

A simplifying assumption I make is that police force size does not change in response to arrest numbers (no reverse causation). As a simplification to the regression calculations, I use notions of size calculated with data from the first month of 2013. Thus, it is imperative that this assumption be close to true so that the assumed values of the variables are still valid in the later months of the study.

Related to the previous assumption, it is also important the municipality populations don't change in response to arrest numbers, since one view of force size incorporates the municipality population. Since I study a relatively short time-frame of five years — not long enough for citizens to move in or out in response to crime rates — this appears to be a reasonable assumption.

I then break the overall analysis into three sub-stages, outlined below.

3.1 First stage - No Time Horizon

In the first stage, I investigate the effect of a police killing only on the arrests in that same month. I seek to estimate

$$Y_{it} = \beta_0 + \beta_1 \text{shoot}_{it} + \beta_2(\text{shoot}_{it} \times \text{killed}_{it}) + I + T + \varepsilon_{it}.$$

A traditional difference-in-differences formulation might look like

$$\beta_0 + \beta_1 \text{shoot}_{it} + \beta_2 \text{Post}_{it} + \beta_3(\text{Treat}_{it} \times \text{Post}_{it}) + I + T + \varepsilon_{it},$$

however, the Post column and the "shoot" column are equal since we consider the Post column to be simply an indicator for time $t = 0$ in the no-time-horizon case, so the traditional formulation and my estimating equation are equivalent.

Here, I and T are year and municipality fixed effects to account for intrinsic differences in explanatory or response variables by municipality or year.

The results of this first regression are shown below:

Table 2: Impact of LODD on arrests, same month

	Response Variable log(Total Arrests)
Shoot_{it}	0.0512***
	(23.33)
$\text{Shoot}_{it} \times \text{Killed}_{it}$	0.0289
	(0.83)
Constant	2.180***
	(2447.66)
Observations	330456

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

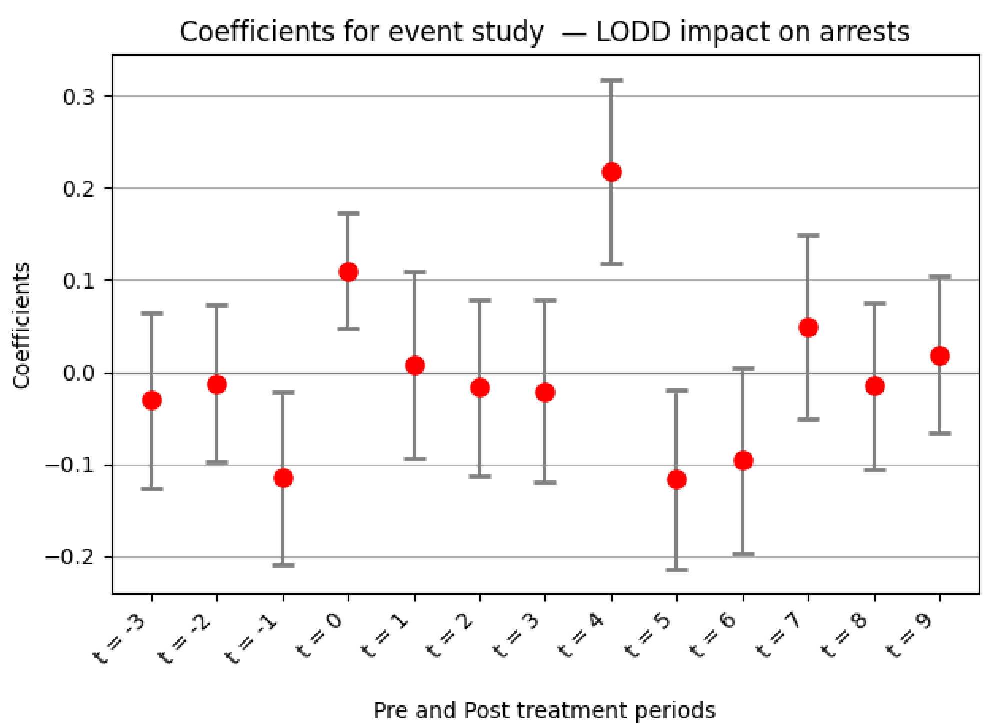
The main coefficient of interest from this regression is an insignificant $\beta_2 = 0.0289$. However, the table as a whole suggests that even an officer's being *shot* in the line of duty has a significant impact on total arrests in the municipality (in particular, prompting a percent increase of 5.07 in arrest numbers), though additionally being killed may not make as much of an impact.

3.2 Second stage - Event Study

In the second stage, I add post treatment indicators for 1, 2, . . . , 9 months after a police shooting and killing since it is possible that the effect of an officer death is not restricted to the month of treatment. In this stage, if there was a shooting in month t in municipality i, the dummy Post_k takes on value 1 in month $t+k$ in municipality i; I also add pre-treatment indicators to confirm the absence of pre-trends or an anticipation-like effect. Note that the interaction term coefficients represent the effects relative to untreated (shooting but no killing) baselines, and are thus the coefficients of interest. Taken together, the estimating equation becomes

$$\begin{aligned} Y_{it} = \beta_0 &+ \sum_{j=-3}^{9} \beta_{\text{Post, j}} \text{Post}_{j,it} + \sum_{j=-3}^{9} \beta_{PT,j} (\text{killed}_{it} \times \text{Post}_{j,it}) \\ &+ I + T + \varepsilon_{it}. \end{aligned}$$

The results of this event study are shown below, with bars indicating confidence intervals for coefficient values; the complete table of values, and graph of $\beta_{\text{Post},j}$ baseline coefficients is given in the appendix. The coefficients are not normalized in this graph.



Firstly, the graph confirms an absence of pre-trends, since there are no patterns in pre-treatment coefficients; this affirms the validity of the parallel trends assumption, as there appears to be no systemic pattern of difference in the response variable among the treated municipalities prior to the introduction of the treatment. There is a significant coefficient at $t = -1$, but this can be reasonably attributed to random variation. The apparent lack of pre-trends is also a positive sign for the robustness of the treatment effect estimation, as pre-trends may bias the coefficient estimates in the post-treatment months, though it is not a guarantee of unbiasedness. Additionally, it suggests that the treatment and control groups were similar with respect to other unobserved factors, reducing the risk of the estimated coefficients' being affected by treatment-correlated unobserved variables.

Regarding the actual treatment effect, I find significance in month 0: arrests are recorded at the end of the month, so the response reflects arrests within 4 weeks of the treatment. There is a positive effect in this immediate aftermath, suggesting an uptick in arrests reflective of the hyper-vigilance theory — the coefficient $\beta_{PT,0} = 0.110$ suggests an 11.6 percent increase in arrests in month $t = 0$. I would hope that the uptick in arrests does not occur before the treatment — due to the absence of positive coefficients in the months prior to treatment, this seems to be a reasonable assumption: there appears to be no anticipation-like effect.

There is little significance in the nine post-treatment months. There is also no discernible pattern in the signs or magnitudes, so the significant coefficient in month 4 is likely to be random. This lack of pattern also makes it unlikely that there could be realized effects beyond nine months. While [3] finds significant negative effects in months 6-12 beyond the treatment, my lack of effects could be due in part to the sample restriction's removing places with profiles of little-to-no crime, which if included (as they are in [3]) would magnify the effect of the treatment.

3.3 Third stage - Triple Difference

Finally, I add the size of the police force as a regressor, estimating a triple difference equation with the following interaction terms:

$$\begin{aligned} Y_{it} = \beta_0 + \beta_1 \text{Post}_{it} + \beta_2(\text{killed}_{it} \times \text{Post}_{it}) + \beta_3(\text{ForceSize}_{it} \times \text{Post}_{it}) + \\ \beta_4(\text{killed}_{it} \times \text{ForceSize}_{it} \times \text{Post}_{it}) + I + T + \varepsilon_{it}. \end{aligned}$$

The Post_{it} variable here is a consolidated Post indicator variable, which is assigned a 1 if any of the Post_i variables from the previous stage were set to 1. I use the same window as the previous stages. The coefficients capture an average treatment effect.

I run this regression twice, with two differently calculated force size metrics; each represents a different interpretation that might affect the response differently.

The first is a measure of size relative to other police force sizes, which can alternatively be thought of as a measure of "absolute size." This can be interpreted as a measure of familiarity or fraternity within the police force. I predict that a smaller police force according to this metric will be tied to higher-magnitude coefficients since the impact of a death might be felt more strongly in more absolutely small police forces.

The second metric is a measure of force size relative to municipality population. This can be interpreted as an indicator of how much cities are willing to invest in their police forces or the value of a police force to a city. Cities with a high ratio of force size to population might be more demonstrably supportive of the police or might offer other support resources to the police in the aftermath of police killings. I would expect that by this metric also, a larger force size is tied to smaller-magnitude coefficients. I specify only a magnitude and not a sign since the sign of the treatment effect is unclear.

The results of the regressions are below:

Table 3: Impact of LODD and "absolute" force size on arrests

	Response Variable log(Total Arrests)
Post-Treatment$_{it}$ (PT$_{it}$)	0.00761 (1.70)
PT$_{it}$ × Killed$_{it}$	0.178* (2.13)
PT$_{it}$ × Force Size$_{it}$	-0.0100 (-1.65)
PT$_{it}$ × Force Size$_{it}$ × Killed$_{it}$	-0.169 (-1.88)
Constant	2.457*** (823.10)

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Impact of LODD and population-relative force size on arrests

	Response Variable log(Total Arrests)
Post-Treatment$_{it}$ (PT$_{it}$)	-0.00257 (-0.61))
PT$_{it}$ × Killed$_{it}$	0.0826 (1.83))
PT$_{it}$ × Force Size$_{it}$	0.00981 (1.62))
PT$_{it}$ × Force Size$_{it}$ × Killed$_{it}$	-0.0962 (-1.55))
Constant	2.455*** (861.14)

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

I again find no significant effect due to either force size in the treated cases, or the treatment itself. Interestingly, when comparing the magnitudes of the coefficients in each notion of force size, it seems that there is a more significant reduction in arrest numbers when considering absolute force size, giving credence to the theory that familiarity with the officers may be the more important driver compared to resources in the police departments.

Another observation is that the $\text{PT}_i \times \text{Killed}_{it}$ term, which was the coefficient of interest in the event study analysis, is significant in this analysis in the absolute force size case. This is possible if police force size is correlated with the treatment effect in a way that is not reflect in the original event study.

3.4 A comment about the null effects

Most computed coefficients in the study appear to be nulls — there are few β significantly different from 0. The precision of this result is difficult to assess with certainty, but some contributing factors are the the imbalance of the dataset, large dataset size, and variability in the dataset.

The imbalance in the dataset — that is, the very small ratio of positive examples to total examples — makes it likely that the results are noisy nulls. In each year of data, there were around 40 - 50 officers killed, constituting a very small percent of the examples in the 10,000 municipalities and 12 months per municipality of data. This lack of positive examples could be resulting in high variability and standard errors in the estimates; it could also result in increased likelihood of a Type 2 error, underestimating the significance of the results. While in general, large dataset size increases precision, in this case, the large dataset size when taken together with the small number of positive examples may increase the noise in the result. The effect of this imbalance is present across all regressions and results.

4 Conclusions and Future Work

This paper contributes to work in policing and studying drivers of police behavioral response. I provide new estimates with a sample restriction for the effect of a line of duty death on arrest numbers in the relevant municipality; I also introduce a heterogeneity, police force size, and study its effect when combined with the line of duty death treatment. I find a significant positive treatment effect in the exact month of treatment, and no significant effects in the following months or with the addition of the force size. There is a possible significant effect of the original treatment in the triple-difference case, which could be due to unobserved correlation between the treatment effect and the force size. I find that familiarity with the fallen officer is a bigger driver of downstream behavior change than perhaps the resourcedness of the police department. While the results were found to be non-significant, the relatively small number of officers killed in the dataset contributes to a possibly noisy null, and may result in underestimations of the significance of the treatment effect.

Additional work on the possible effect of the arrests or policing behavior on the population should investigate potential reverse causality. It could also be worth investigating the types of arrest stops made to lend credence to the $t = 0$ hypervigilance theory suggested by the coefficients — the number of stops for trivial offences' increasing would affirm that officers act hypervigilant. I would also like to re-impose the sample restriction via matching: using a model to identify municipalities with high propensity for assaulted police officers could be another way to restrict the sample to municipalities with crime profiles similar to those where officers might be killed.

References

[1] Highest to lowest - prison population rate, 2023.

[2] Amelia Cheatham and Lindsay Maizland. How police compare in different democracies, 2022.

[3] Sungwoo Cho, Felipe Goncalves, and Emily Weisburst. The impact of fear on police behavior and public safety. *NBER Working Paper 31392*, 2023.

[4] Oeindrila Dube, Sandy Jo MacArthur, and Anuj K. Shah. A cognitive view of policing. *NBER Working Paper 31651*, 2023.

[5] Will Johnson. Community policing: Much more than walking a beat, 2018.

[6] Jacob Kaplan. Uniform crime reporting program data: Law enforcement officers killed and assaulted (leoka) 1960-2021, 2020.

[7] Jacob Kaplan. Uniform crime reporting (ucr) program data: Arrests by age, sex, and race, 1974-2021, 2020.

[8] Laurence Miller. Line-of-duty death: Psychological treatment of traumatic bereavement in law enforcement. *International Journal of Emergency Mental Health*, 9:13–23, 2007.

[9] Andreas Olden and Jarle Moen. The triple difference estimator.

Appendix

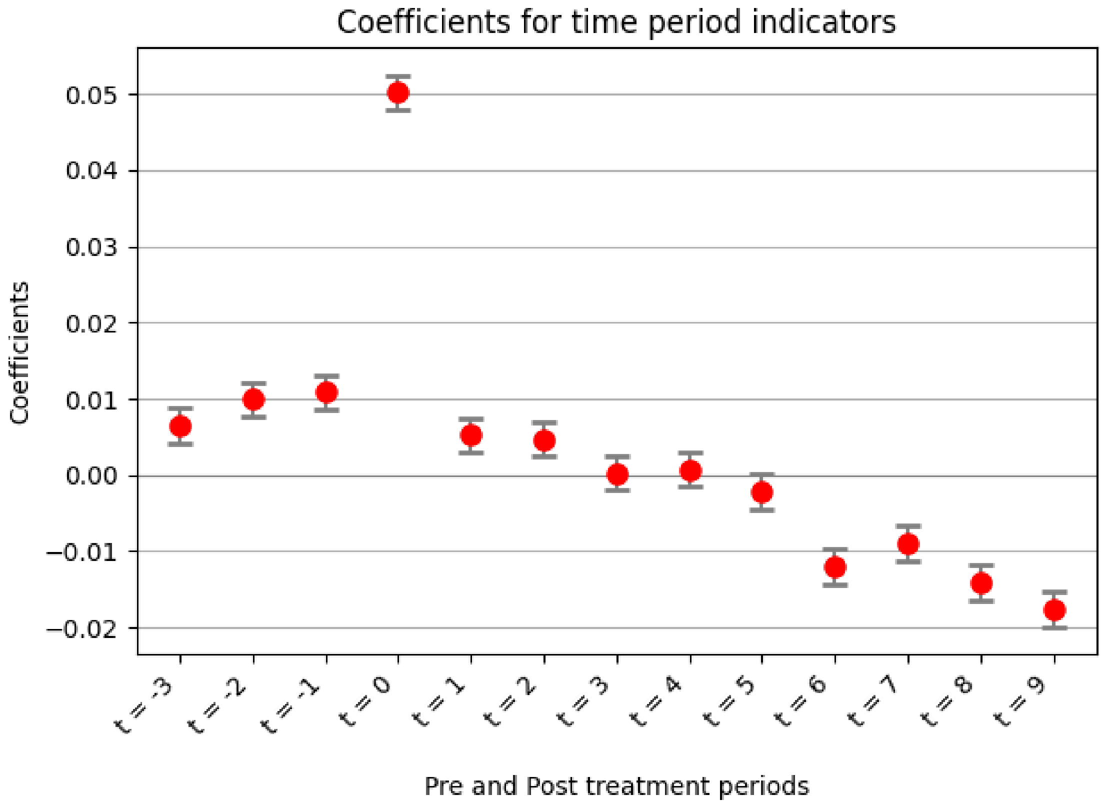
Here is the table of baseline time trajectory coefficients from the second-stage event study, along with a plot of the coefficients — these are the coefficients denoted $\beta_{\text{Post},j}$ in the estimating equation.

	log(total arrests)	t-stat	p $\geq$ | t |	95-percent confidence interval
t = -3 (Post$_{-3,it}$)	0.0065** (.00224)	2.90	0.004	[.0021, .0109]
t = -2	0.0100*** (.00224)	4.44	0.000	[.0056, .0143]
t = -1	0.0109*** (.00224)	4.87	0.000	[.0065, .01527]
t = 0	0.0503*** (.00221)	22.77	0.000	[.0459, .0546]
t = 1	0.0053* (.00225)	2.35	0.019	[.0009, .0097]
t = 2	0.0047* (.00226)	2.08	0.038	[.0003, .0091]
t = 3	0.0003 (.00227)	0.11	0.910	[-.0042, .0047]
t = 4	0.0007* (.00229)	0.31	0.757	[-.0038, .0052]
t = 5	-0.0022 (.00231)	-0.93	0.351	[-.0067, .0024]
t = 6	-0.0120*** (.00232)	-5.17	0.000	[-.0166, -.0075]
t = 7	-0.0090*** (.00234)	-3.85	0.000	[-.0136, -.0044]
t = 8	-0.0141*** (.00235)	-6.01	0.000	[-.0187, -.0095]
t = 9	-0.0176*** (.00235)	-7.47	0.000	[-.0221, -.0129]
Observations	330456			

95% confidence intervals in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The corresponding plot of coefficients, which reflect the baseline effect of a shooting for many months, is shown. Coefficients have not been normalized:



Like the single coefficient from the first stage, we see that the effect of a shooting itself is significant in many months post-shooting. The effect in month $t = 0$, which is captured between 0 and 4 weeks post-shooting, appears most significant with a strong positive effect on arrests; in the remaining months, there is a clear downward trend in the impact of the shooting, starting with smaller positive effects for up to 3 months post shooting, and ending with significant negative effects in months 6 to 9.

Here is the table of interaction term coefficients from the second stage event study; these are the coefficients corresponding to the graph shown in the paper.

	log(total arrests)	t-stat	p $\geq$ | t |	95-percent confidence interval
t = -3 (killed$_{it}$ $\times$ Post$_{-3,it}$)	-0.0307 (.095)	-0.32	0.748	[-0.218,0.156]
t = -2	-0.0125 (.086)	-0.15	0.884	[-0.180,0.155]
t = -1	-0.115 (.093)	-1.24	0.216	[-0.298,0.0674]
t = 0	0.110 (.063)	1.74	0.082	[-0.0140,0.233]
t = 1	0.00746 (.102)	0.07	0.942	[-0.192,0.207]
t = 2	-0.0167 (.095)	-0.18	0.861	[-0.204,0.170]
t = 3	-0.0209 (.099)	-0.21	0.833	[-0.215,0.173]
t = 4	0.218* (.100)	2.18	0.029	[0.0219,0.414]
t = 5	-0.116 (.097)	-1.20	0.231	[-0.307,0.0741]
t = 6	-0.0957 (.101)	-0.95	0.343	[-0.293,0.102]
t = 7	0.0490 (.100)	0.49	0.624	[-0.147,0.245]
t = 8	-0.0151 (.090)	-0.17	0.867	[-0.192,0.162]
t = 9	0.0190 (.085)	0.22	0.823	[-0.147,0.185]
Constant	2.182*** (.002)	1449.31	0.000	[2.179,2.185]

95% confidence intervals in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All code for this paper can be found at https://github.com/NSrinidhi/14.33-policing-project.