An evaluation of Crisis-Intervention Team (CIT) training

Nemschoff, Danielle

University of Chicago

12 October 2022

An Evaluation of Crisis Intervention Team (CIT) Training

Danielle Nemschoff*

October 12, 2022

Police officers in the United States are most often the first responders to a mental health crisis. The most popular training method for these responses among US police departments is crisis-intervention team (CIT) training. This paper provides the first estimates of the causal effect of CIT training on a police officer's propensity to use force and make an arrest. I implement a difference-in-differences framework using future trainees as controls to compare officer use of force and arrest of trained officers to those of untrained officers. I do not find a statistically significant effect of CIT training on either use of force or propensity to arrest.

1 Introduction

Mental health is a growing crisis in the United States and has been gaining both media and political attention in the wake of the COVID-19 pandemic and the continued epidemic of mass shootings. Specifically, there has been a growing discourse on the role of police, and the criminal justice system as a whole, in mental health crisis intervention. Currently, Los Angeles County Jail, Cook County Jail, and Riker's Island Jail Complex each hold more mentally ill inmates than any remaining psychiatric hospital in the United States (Treatment Advocacy Center, 2016) and every one of these individuals must have interacted with the police at least once. Around 21 to 38% of 911 calls nationally are estimated to be related to mental health (Treatment Advocacy Center, 2016). Not only do police frequently interact with those in mental health crisis, but these interactions are often fatal. Individuals experiencing a mental health crisis are 16 times more likely to be killed during an interaction with police (American Journal of Preventative Medicine). Since 2015 22% of those shot and killed by police were having a mental health crisis (The Washington Post, 2022).

In response to the growing demand for police response to mental health crises many departments have been implementing crisis-intervention team (CIT) training programs. The goal of CIT training is to teach police officers how to de-escalate mental health crises and divert civilians with a mental health issue from the criminal justice system to health care and social services. To my knowledge, this is the first paper to empirically study the causal effects of CIT training on police use of force and propensity to arrest. I use an extensive data set from New Orleans covering every police call for service and every self-initiated police stop from 2017-2020 to assess the impact of CIT training on police use of force and rates of arrest. In my preferred specification I implement a difference-in-differences model using future trainees as controls for officers already trained in each month. I find no evidence CIT training reduces officer use of force or propensity to make an arrest when responding to a mental health calls. I also check for spillover effects of the training on either types of calls officers respond to and find no effect on use of force or arrests for non-mental health incidents, drug-related incidents, or violent crime.

This paper contributes to the small but growing economics literature on police training. McLean et al., 2020 studies the impact of social interaction training for police and similarly find it has no impact on officer use of force. Rather than evaluating a specific program Adger et al., 2022 attempt to identify how the behavior of a field training officer impacts their trainees and find that recruits who had field training officers with higher use of force also demonstrate a higher use of force, shedding light on a possible avenue for effective change. Owens et al., 2018 and Dube et al., 2022 both develop, implement and evaluate training

*I would like to thank my two readers Dan Black and Manasi Desphande for their helpful comments. I would also like to thank Bocar Ba for his support and Derek Neal for constructive feedback. Thank you all so much.

programs and find that officers who receive training were less likely to use force and less likely to make discretionary arrests. The training in Owens et al., 2018 focuses on procedural justice while the training in Dube et al., 2022 is cognitive-based. Their results may speak to the importance of the modality of training.

CIT training is only one method for emergency response to mental health crises. Another proposed method is one of community response, where mental health service provideers respond to calls for service without police. Dee and Pyne (2022) study a community response to mental health crises in Denver and found the program reduced reports of less serious crimes by 34%. From a policy perspective their work complements this paper by providing a possible alternative emergency response program for mental health crises since I am unable to find clear evidence that one of the more common police-based programs is effective.

The paper proceeds as follows: Section 2 provides a background on police procedure in New Orleans and provides a description of crisis-intervention training. Section 3 discusses the data sets used, data construction and summary statistics. Section 4 explains my empirical strategy, Section 5 presents results, and Section 6 concludes.

2 Background

2.1 Background on 911 calls and police response process

A call for service begins either by a 911 call or an officer-initiated event. In the case of a 911 call, the call is taken at the call center in New Orleans. The call's metadata (location and time) are recorded in the Computer-Aided Dispatch System (CAD). The 911 call-taker classifies the type of incident and records this in CAD. They then triage the call, decide which service response (if any) is needed, assign a priority level and dispatch the appropriate service. If the call is deemed to require a police response, officers are dispatched in one-man units by priority level of the call and proximity to the incident[1]. The responding officers dispatch time and arrival time are also recorded in CAD.

The responding officers assess the incident and then enter an incident classification and priority level in CAD. These can be different than the original classification and priority assigned by the dispatcher. Once the responding officers arrive on scene they assign the call a disposition based on what how the incident is resolved. The call disposition may be deemed *null* or *void* if it is a false call. There may have been a valid incident but the concerned parties be gone by the time officers arrive, in which case the call disposition is classified as *gone on arrival*. If there are subjects still present upon arrival the call disposition will either be *necessary action taken (NAT)* or *report to follow (RTF)*. An incident receives a disposition of NAT if no arrests are made, no police reports are filed, and no substantial action is taken by the police at the scene. I do not observe anything further about the incident or officers if the call is deemed NAT. If the disposition is recorded as *RTF* the responding officers also write a police report for the incident, recording number of offenders, number of victims and demographics on both the offenders and victims if known. I observe a police report for 20-30% of calls a year. The officers use their own observations, victim interviews and witness statements to fill out these reports.

Per New Orleans Police Department (NOPD) policy if any force is deployed beyond non-resisted handcuffing, a use of force report must be written. This is written by the supervising officer at the scene and should include demographics on each officer and subject involved, record the type and level of force used by each officer, injury status of officers and subjects involved, and any influencing factors.

[1] According to my correspondence with NOPD Public Information, the default approach is to dispatch one officer per call. However, the dispatcher makes this decision based on the priority of the call, the description of the call, and officer availability and can decide to dispatch as many one-man units as is deemed appropriate. The responding officers will be those available and closest proximity to the incident. If an officer calls for backup all available units respond until a dispatcher or supervisor tells them the situation is under control.

2.2 Crisis-Intervention Team Training

A very popular model for helping first responders adapt to an increasing number of interactions with those who may be mentally unstable is crisis-intervention team (CIT) training. The version of CIT used in most police departments today is called the Memphis Model; it was developed by the National Alliance on Mental Illness (NAMI) of Tennessee after the high-profile incident of a man in a mental health crisis in 1988 (The University of Memphis CIT Center, 2022). Since then, as more media and popular culture attention has been drawn to policing in general, and issues of mental health in particular, many more police departments have implemented CIT programs. In 2008 NAMI estimated that there were about 400 CIT programs across the US and by 2019 they estimated there were over 2,700 programs (National Alliance on Mental Illness, 2022).

While there has been little academic study of CIT, there are plenty of anecdotal claims success. On their website NAMI claims "...programs like CIT reduce arrests of people with mental illness while simultaneously increasing the likelihood that individuals will receive mental health services." They also make the specific claim that "in Memphis, for example, CIT resulted in an 80% reduction of officer injuries during mental health crisis calls. " A unique advantage of my data is that I am able to view rates of arrest and use of force, and so can analyse the validity of each of these claims directly.

CIT in New Orleans

In 2008 an NOPD officer was killed by a man suffering from mental illness with her own gun (Reckdahl, 2021), prompting the commitment to establish a CIT program and train at least 20% of the patrol officers by August 2016 in the federal consent decree in 2013. As of 2021 there had been 17 cohorts of officers to receive CIT training (1 class in 2015, 3 in 2016, 4 in 2017, 2 in 2018, 2 in 2019, 1 in 2020, and 1 in 2021). In total 365 officers have gone through training representing around 30% percent of all patrol officers[2].

Calls deemed appropriate for CIT trained officer response are those the dispatcher deems to be a "mental patient", "suicide threat", "suicide attempt" or "suicide"[3]. CIT trained officers responded to about 60% of the calls with these signals over the 4 years of data. Protocol is for the dispatcher to locate the nearest CIT trained officer when a mental health call comes in and dispatch them if they are available. If a CIT officer is not available, the closest officer is dispatched. Per the NOPD Operations Manual: "The ideal resolution for a crisis incident is that the individual is diverted from the criminal justice system and connected with resources that can provide long-term stabilizing support."

3 Data

3.1 Data sets

In 2010 the Department of Justice started an investigation of the New Orleans Police Department (NOPD) in which resulted in a consent decree in 2013. One of their mind findings was a pattern of excessive force and a practice of inadequate reporting on use of force incidents. In light of this, NOPD was required to collect and report use of force data, which they started publishing in 2016. They additionally started publishing calls for service and electronic police reports in 2017. The data I use is a mix of publicly available data and data made available by FOIA requests. For the years 2017-2020 I have data at the incident level of every call for service made (includes police initiated responses as well as 911 calls), electronic copies of police reports, a list of every call CIT officers responded to and a report on each incident of use of force. I also have the demographics of all commissioned officers in each district over these 4 years, as well as information on which officers went through CIT. The novelty of this data set is that I can track every arrest and use of force to its originating call for service. Additionally, dispatch data allows me to circumvent the usual issue of selection

[2]From communication with NOPD, I learned that the majority of officers signed up voluntarily while some were selected from the program. There is no documentation available on which officers volunteered versus were selected.

[3]breakdown is included in the Appendix.

in policing data. Most police research uses stop & search data, vehicle stop data, or police reports. This data is necessarily conditional on an officer initiating contact. Police officers exercise substantial discretion in who they approach and how they respond and so we'd expect this data to suffer from selection bias. With dispatch data I can see every call any officer was sent to, whether or not there was a report written, arrest made, or force used.

Call data: includes the call's priority level assigned by the dispatcher, a classification of the type of call (assault, homicide, disturbance, etc.), date and time call was received, time officers were dispatched, location at the block-level, and the employee ID of officers dispatched. About 29% of calls from 2017-2021 are missing the employee ID numbers.

Police report data: includes the incident's priority level, a classification of the type of incident, date and time of incident, location at the block-level, race and gender of each offender as well as the race and gender of each victim. Each report also includes the officer who wrote it. It is important to note the police reports are written by the responding officers and so is from their point of view. Importantly, any arrest has an associated police report and so this is where I obtain my arrest data.

Use of force data[4]: includes date and shift of the incident (an 8 hour window); location by police district; name, race, gender, age, and years of service of each police officer involved; gender, age, and race of each subject involved; type and level of force used; injury status of officer and subject; and whether subject was hospitalized. NOPD provided me with a crosswalk to directly link each Use of Force incident back to its originating call for service.

CIT data: a list of incident numbers to which CIT officers were dispatched and lists of each class of officers trained with their training completion date [5].

Officer demographic data: includes the name, race, gender, age, assigned police district, hire date and retire date (if applicable) of every commissioned NOPD officer for each year 2017-2020.

3.2 Descriptive Statistics

Data Construction: I first exclude police-initiated events from the sample to avoid selection issues. I then restrict the sample to calls with a disposition of "report to follow" or "necessary action taken" as these are the calls with a likely police-civilian interaction (this excludes calls labelled "gone on arrival", "unfounded", and "null"). This leaves me with just under 1 million incidents over just more than 3 years. Mental health calls are defined as calls designated "mental patient", "suicide", "suicide threat", or "attempted suicide" by the dispatcher[6].

Table 1 reports observation counts in the data. Less than 1% of the incidents captured in these reports involve reported use of force, around 2% of calls are mental health related a year, and between 14-18% of incidents result in arrest per year. CIT trained officers responded to around half of the calls for service each year.

[4] A full definition of use of force is provided in the appendix.

[5] Training began at the end of 2015 but NOPD did not begin collecting data on dispatch until 2017 so I know who was trained prior to 2017 but I don't know which calls they were dispatched to.

[6] Suicides make up less than 0.05% of CIT calls and so are not driving the results although they may be thought of as different since they may not immediately endanger the responders' lives.

Table 1: Calls for Service

	2017	2018	2019	2020*
# calls	324,284	308,861	249,3776	25,378
% mental health	0.02	0.02	0.02	0.02
% arrest	0.17	0.18	0.18	0.14
% force	0.001	0.001	0.001	0.001
% CIT	0.55	0.56	0.56	0.14

Notes: *Only January - March 2020. Row 1 is the number of calls for service per year with a disposition of necessary action taken or report to follow.

Table 2 reports the number of instances of use of force per year and the percentage of use of force cases that were mental health calls. Mental health calls have a disproportionately high-share of use of force incidents given their share of overall calls. Mental health calls are seven times more likely to result in force than the average call. The top three call types to result in force in order are "other", "drugs" and "mental patient".

Table 2: Use of Force

	2017	2018	2019	2020*
# instances	351	278	253	24
% mental health	0.08	0.12	0.14	0.01

Note: *Only January - March 2020. Universe is all valid calls for service per year. Row 2 is percent of force incidents that were mental health calls.

There are eight police districts in New Orleans. The distribution of officer demographics is relatively steady over-time but does meaningfully differ between districts. Years of service and office age are balanced between districts but racial and gender composition are not. For instance, District 4 has a much higher share of Black and female officers compared to any other district. Average demographics of each district are presented below.

Table 3: Demographics by District

district:	1	2	3	4	5	6	7	8
# officers	80	84	80	78	82	76	87	103
Avg Age	38	40	40	41	39	38	38	40
Avg Years of Service	11	12	12	13	13	12	12	13
% Black	0.38	0.47	0.34	0.71	0.49	0.40	0.54	0.43
% white	0.42	0.42	0.54	0.21	0.31	0.47	0.24	0.41
% female	0.17	0.23	0.17	0.31	0.25	0.21	0.26	0.19

Note: Statistics averaged over 2017-2020.

3.3 CIT

CIT trained officers are relatively uniformly distributed across districts and years (ranging from 5-12% of officers). The gender and racial distribution of CIT trained officers is in align with the overall distribution of all commissioned officers. Because CIT training is mostly voluntary it is important to look at patterns of force use between officers who chose to participate versus those who did not. Because data collection on use of force did not begin until 2016 and the first CIT training cohort began in 2016, I am not able to look at pure pre-trends between ever-treated and never-treated officers. Instead, I compare use of force for officers trained in CIT after 2016 *from before they are trained* to the use of force of never trained officers. On average officers who will receive CIT training used force 20% more and made an arrest 4% more often than officers who are never trained. This suggests those with a higher propensity to use force may be the ones selecting into training.

Force: Even after CIT training began, compared to any other type of call, mental health calls are significantly more likely to end in force. Furthermore, mental health calls with a CIT officer responding end in force 6.5x more often than mental health calls *without* a CIT officer responding. In 2020 mental health calls with a CIT response ended in force 50% more often than those without a CIT trained officer responding.

Table 4: % of Mental Health Incidents that Result in Force by CIT Status

	CIT response	non-CIT response
2017	0.0090	0.0012
2018	0.0074	0.0060
2019	0.0099	0.0060
2020	0.0099	0.0066

Note: Col 1 is the percent of mental health calls that had a CIT officer respond that has reported use of force. Col 2 is the percent of non-mental health calls that had a CIT officer respond that has reported use of force.

Table 5: Officer Demographics

	CIT trained	untrained	overall
% female	0.30	0.20	0.23
% black	0.47	0.50	0.50
avg age	38.4	42.9	41.2
avg years of experience	9.61	12.6	11.6

Note: Col 1 is the percent of CIT trained officers who are female and Black, their mean age and mean years of experience. Col 2 is the percent of non-CIT trained officers who are female and Black, their mean age and mean years of experience. Col 3 presents the average demographics for all officers.

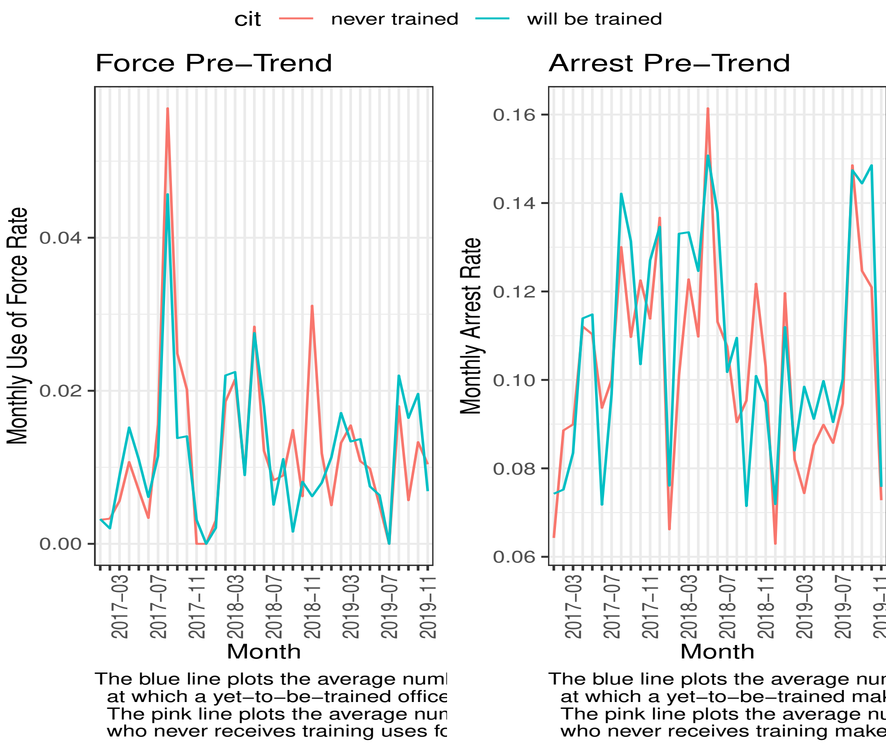
4 Empirical Strategy

I consider a staggered adoption design with N officers over T months. $i \in \{1, ..., N\}$ indexes the officers and $t \in \{1, ..., T\}$ indexes the months. Officers can receive training in any time period or not at all. The month officer i completes training is considered the month of treatment. Once an officer is trained, they are considered treated for all periods afterwards. For each officer in each month there exists the set of potential outcomes $\{Y_{0it}, Y_{1it}\}$. The key object of interest is $Y_{1it} - Y_{0it}$, the difference in an officer's propensity to use force/make an arrest before and after training. The ideal counterfactual for a trained officer would be the potential outcome Y_{0it}, the outcome for this same officer i if they had not received training. This is inherently unknowable. The various estimation techniques I employ are at a basic level attempting to find a suitable counterfactual to credibly estimate a causal effect.

The naive way to estimate $Y_{1i} - Y_{0i}$ would be to compare the use of force and arrest rates of trained officers post-training with the use of force and arrest rates of untrained officers. To interpret this difference as the causal impact of CIT training we would have to believe that trained and untrained officers do not differ in anyway other than training. However, as demonstrated above officers who attend training systematically differ from those who do not. Instead, we can use officer i *before* training as the counterfactual for officer i *after* training. This gives us the classic difference-in-differences estimator I use as my base empirical specification.

This naive estimator implicitly includes both individual officer effects and month effects in the error term. A natural first step would be to implement a two-way fixed effects model, which explicitly controls for unobserved individual officer effects and unobserved calendar month effects, improving the estimate of the treatment effect relative to the naive estimator described above. However, this empirical strategy relies on the assumption of parallel pre-trends, which is not satisfied as demonstrated in Figure 1.[7]

Figure 1: Pre-Trends



Additionally, while trained and never-trained officers do appear similar on observable characteristics, we may be worried about unobservable differences. To test whether training predicts use of force or propensity to make an arrest I limit the sample to incidents involving officers who are never-trained and officers who will be trained *prior* to the officer receiving training. I regress an indicator for force use/arrest on an indicator for will-be-trained and controls for call characteristics. I find that training is significantly positively related to both force use and arrest[8]. This indicates that never-trained officers systematically differ from trained officers and so the never-trained officers may not be an appropriate counterfactual for trained officers. This

[7]The static specification where all treatment is pooled into one post-period and a dynamic version that allows for heterogeneity across training cohorts are presented in the Appendix.

[8]Appendix Table 13

motivates my estimation technique of dropping the never-treated and using not-yet-treated officers as controls.

To address the potential violation of the parallel trends assumption discussed above I estimate a difference-in-differences style equation using officers who will receive CIT training but have not *yet* been trained at time t as the control group for officers trained at time t. This approach resembles the empirical strategy adopted by Desphande & Li (2019), in which they exploit variation in the timing of treatment. For each month t I label officers who had finished training by that month as treated and use officers who are trained in the next calendar year or later as the control group in month t. In practice, I create a separate data set for each training cohort and stack them together. For example the data set for the training cohort February 23, 2018 would include all officers who completed training on this day, marked as treated, as well as all officers who completed training on any date a year or more later (after February 23, 2018). This second group is the control group. This means that individual officers can be included as both treated and controlled units in the final stacked data set. This approach exploits plausible exogeneity in the *timing* of training. To support the validity of this assumption I limit my sample to only officers who will be trained and test whether individual training cohorts differ on demographics or pre-training use of force and propensity to arrest[9]. I look at the percentage of officers who are Black, white, and male in each cohort, the average number of years of service when trained and the mean percent of monthly calls to end in force/arrest for each training cohort. I find no systematic pattern across cohorts, further supporting that conditional on being trained timing of training is as-good-as random.

Using officers who will receive CIT training but have not *yet* been trained at time t as the control group I then estimate the following equation:

$$Y_{igt} = \alpha_i + \gamma_t + \beta_0 Treated_{ig} + \sum_k D^k_{gt} + \sum_k \beta_k \left(Treated_{ig} \times D^k_{gt}\right) + \epsilon_{igt} \tag{1}$$

Y_{igt} is an outcome for officer i from training group g in month t. α_i are officer fixed effects and γ_t are month fixed effects. $Treated_{ig}$ is an indicator equal to 1 if officer i is trained in group g. D^k_{gt} is an indicator equal to 1 if month t is k months after (or before) the month of training. The month of training is zeroed out since training date within month varies and it is ambiguous whether an officer is treated or not in this month. The sample is all mental health incidents responded to by an officer who receives CIT training during the window of observation.

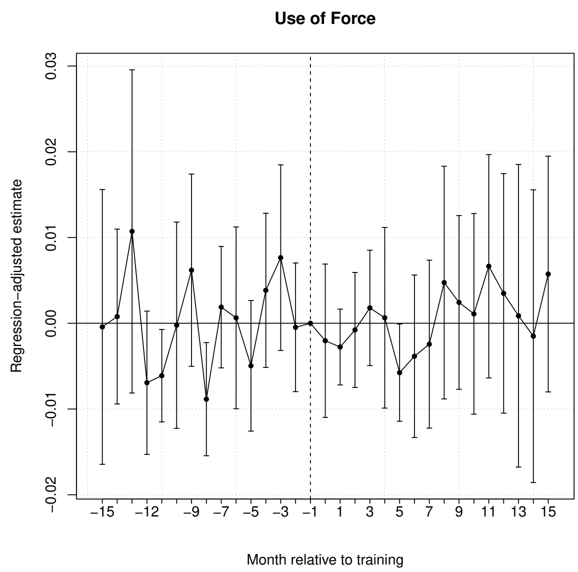
5 Results

I estimate the impact of CIT training on officer use of force and propensity to arrest by dropping the never-treated officers and using officers who *will be treated* as the control group for those officers trained in a given cohort. I do this to address the concern of weak parallel trends between treated and never-treated officers. The impact of CIT training on officer use of force and propensity to arrest is indistinguishable from zero. Figure 2 and Figure 3 present estimates of the effect of CIT training on an officer's propensity to use force and to make an arrest respectively. These are estimates of the β_k coefficients from Equation (1). The estimates presented in Table 9 are for the coefficient δ from the equation:

$$Y_{igt} = \alpha_i + \gamma_t + \delta_0 Treated_{ig} + \sum_k D^K_{gt} + \beta \left(Treated_{ig} \times Post_g t\right) + \epsilon_{igt} \tag{2}$$

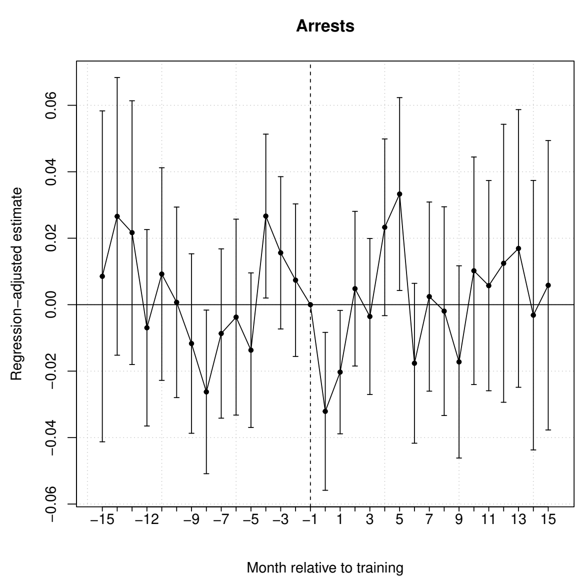
[9] Table 14 and Table 15 in the Appendix.

Figure 2: Effect of CIT Training on Propensity to Use Force



Note: This figure plots estimates of the effect of CIT training on an officer's propensity to use force. These are estimates of coefficients β_k from Equation (1). The dependent variable is a binary indicator for force being used at the incident. The sample is all incidents involving officers who are trained during the period of observation. Controls are yet-to-be-trained officers. The error bars plot the 95% confidence interval of the estimated coefficients.

Figure 3: Effect of CIT Training on Propensity to Arrest



Note: This figure plots estimates of the effect of CIT training on an officer's propensity to make an arrest. These are estimates of coefficients β_k from Equation (1). The dependent variable is a binary indicator for an arrest was made at the incident. The sample is all incidents involving officers who are trained during the period of observation. Controls are yet-to-be-trained officers. The error bars plot the 95% confidence interval of the estimated coefficients.

Table 6: Time-Variation DiD

	Dependent variable:	
	force	arrest
	(1)	(2)
Treated × Post	0.0005	−0.0026
	(0.001)	(0.0031)
Observations	43,324	43,324
R^2	0.026	0.031
F Statistic	1.05^{***}	1.22^{***}
Note:	*p<0.1; **p<0.05; ***p<0.01	

The estimates are for the coefficient δ from Equation (2). Treated = 1 if the officer ever receives training and Post = 1 if the incident occurred after the officer received training. Fixed effects for incident priority, day of week, shift and district are included. The sample is only mental health incidents. Only officers who receive training during the window of observation (2017- Feb 2020) are included. This excludes always-treated and never-treated officers. Standard errors are clustered at the incident level.

5.1 Spillovers

The focus of CIT training is preparing officers to respond to mental health crises, but de-escalation techniques can be easily transferable to other scenarios. While I am not able to find an effect for mental health calls, it is possible officers are using their skills on other calls. To check for possible spillover I use my preferred specification using future trainees as controls and perform the same analysis as above on non-mental health incidents, violent incidents[10] and drug-related incidents[11]. There is no detectable effect of CIT training on use of force or arrest for these incident types. Results are presented below.

[10] I define this as any assault, battery, homicide, weapon possession, or rape.
[11] I define this as any incident labelled a drug violation or drunk.

Table 7: Non-mental health incidents

	Dependent variable:	
	force	arrest
	(1)	(2)
Treated × Post	−0.00005	−0.00078
	(0.0001)	(0.001)
Observations	1,065,319	1,065,319
R^2	0.025	0.037
F Statistic	23.3***	34.1***
Note:	*p<0.1; **p<0.05; ***p<0.01	

The coefficient presented is δ from Eq.(2): Treated = 1 if the officer ever receives training and Post = 1 if the incident occurred after the officer received training. Fixed effects for incident priority, day of week, shift and district are included. The sample is only mental health incidents. Only officers who receive training during the window of observation (2017- Feb 2020) are included. This excludes always-treated and never-treated officers. Standard errors are clustered at the incident level.

Table 8: Violent incidents

	Dependent variable:	
	force	arrest
	(1)	(2)
Treated × Post	−0.0012	0.0019
	(0.0011)	(0.0034)
Observations	76,854	76,854
R^2	0.017	0.071
F Statistic	1.21***	4.48***
Note:	*p<0.1; **p<0.05; ***p<0.01	

The coefficient presented is δ from Eq.(2): Treated = 1 if the officer ever receives training and Post = 1 if the incident occurred after the officer received training. Fixed effects for incident priority, day of week, shift and district are included. The sample is only mental health incidents. Only officers who receive training during the window of observation (2017- Feb 2020) are included. This excludes always-treated and never-treated officers. Standard errors are clustered at the incident level.

Table 9: Drug/Alcohol-related incidents

	Dependent variable:	
	force	arrest
	(1)	(2)
Treated × Post	−0.0029	0.0069
	(0.003)	(0.010)
Observations	3,483	3,483
R^2	0.35	0.48
F Statistic	2.61^{***}	4.16^{***}
Note:	*p<0.1; **p<0.05; ***p<0.01	

The coefficient presented is δ from Eq.(2): Treated = 1 if the officer ever receives training and Post = 1 if the incident occurred after the officer received training. Fixed effects for incident priority, day of week, shift and district are included. The sample is only mental health incidents. Only officers who receive training during the window of observation (2017- Feb 2020) are included. This excludes always-treated and never-treated officers. Standard errors are clustered at the incident level.

5.2 Robustness

A possible concern would be if the number of calls and the composition of type of calls change after officer is trained. It can be highly stressful to be speaking with those in acute mental health crises all day and if trained officers are only assigned to mental health incidents post-training, we could imagine their patience running thin by the end of the day, resulting in potentially worse outcomes. To investigate this I compare the average percent of calls a trained officer responds to per month classified as mental health calls in the months before training (17.8%) to the average percent of calls an officer responded to per month classified as mental health calls in the months after training (21.7%) and am unable to reject the null hypothesis that the sample means are equal.

A concern from the static TWFE model is the causal interpretation of the treatment coefficient. Under the parallel trends assumption, following de Chaisemartin and D'Haultfoeuille (2020):

$$E[\phi] = E\left[\sum W_{g,t} TE_{g,t}\right]$$

where $W_{g,t}$ is the weight assigned for treatment group g at time t and $TE_{g,t}$ is the treatment effect for this group and time period. The decomposition in Goodman-Bacon (2019) shows that this may produce negative weights, especially if there are groups that are treated most or all of the time. Negative weights themselves are not necessarily an issue if there are homogeneous treatment effects, but in the case of heterogeneous effects can bias treatment effects estimates (De Chaisemartin and d'Haultfoeuille 2020, Goodman-Bacon 2018).

However, we still have the negative weights concern. This is directly testable since the weights are proportional to the residuals from a regression of treatment on unit and time fixed effects (De Chaisemartin and d'Haultfoeuille 2020)[12]. Approximately 8% of treated observations receive negative weights and these observations tend to be later months in the observation period for officers trained early[13]. As noted by Jakiela (2021) a sufficiently large never-treated group combined with sufficient pre-treatment data will ensure non-negative weights. Only 30% of officers ever receive training so a large never-treated group exists. To address point two I estimate a version of the two-way fixed effects model where I restrict my sample to

[12] This can be seen by applying the Frish-Waugh-Lovell theorem to a standard DiD model of the form $Y_{it} = \lambda_i + \gamma_t + \beta D_{it} + \epsilon_{it}$.

[13] Residuals are from the regression $Treated_{it} = \lambda_i + \gamma_t + \epsilon_{it}$ where i is an officer and t is a month. λ_i are officer fixed effects and γ_t are month fixed effects.

officers who were never trained or were trained 6 months or more after data collection began, which naturally excludes the always-treated group. Results do not qualitatively change and are presented in the Appendix Table 11.

An additional, and related, concern is that the first CIT training cohort completed training at the beginning of 2016 but data collection did not begin until 2017, therefore I have a group of officers who are "always-treated" across my panel. To test whether always-treated officers differ from the officers whose training I observe I compare use of force and rates of arrest in the post-periods for officers who experience training in my panel and officers who are "always-treated". The plots are presented below. I also test this empirically by regressing outcomes on an indicator equal to 1 if an officer will be treated and is equal to 0 if the officer is "always-treated". The coefficient of interest for both force and arrests is insignificant.

Figure 4: Always Treated vs. Eventually Treated

Note: The blue line plots the average number of calls per calendar month at which an officer who receives training during the observation period use force or makes an arrest after they received training. The pink line plots the average number of calls at which an officer who received training prior to the observed period uses force or makes an arrest in the same calendar month.

6 Conclusion

This paper estimates the impact of crisis-intervention team training on police officers' use of force and propensity to make an arrest using data on all calls for service in New Orleans from 2017 through the beginning of 2020. Results indicate that crisis-intervention training has no discernible impact on officer use of force or propensity to make an arrest. These results help inform the ongoing discussion on alternative emergency response to mental health crises and is a first-step in further work on the role of police as first responders.

Appendix

Use of Force Definitions [14]:

1. Level 1: uses of force include pointing a firearm at a person and hand control or escort techniques (e.g., elbow grip, wrist grip, or shoulder grip) applied as pressure point compliance techniques that are not reasonably expected to cause injury; take-downs that do not result in actual injury or complaint of injury; and use of an impact weapon for non- striking purposes (e.g., prying limbs, moving or controlling a person) that does not result in actual injury or complaint of injury. It does not include escorting, touching, or handcuffing a person with minimal or no resistance.

2. Level 2: uses of force include use of a CEW (including where a CEW is fired at a person but misses); and force that causes or could reasonably be expected to cause an injury greater than transitory pain but does not rise to a Level 3 use of force.

3. Level 3: uses of force include any strike to the head (except for a strike with an impact weapon); use of impact weapons when contact is made (except to the head), regardless of injury; or the destruction of an animal.

4. Level 4: uses of force include all 'serious uses of force' as listed below
   (a) All uses of lethal force by an NOPD officer
   (b) All critical firearm discharges by an NOPD officer
   (c) All uses of force by an NOPD officer resulting in serious physical injury or requiring hospitalization
   (d) All neck holds
   (e) All uses of force by an NOPD officer resulting in a loss of consciousness
   (f) All canine bites;
   (g) More than two applications of a CEW on an individual during a single interaction, regardless of the mode or duration of the application, and whether the applications are by the same or different officers, or CEW application for 15 seconds or longer, whether continuous or consecutive
   (h) Any strike, blow, kick, CEW application, or similar use of force against a handcuffed subject
   (i) Any vehicle pursuit resulting in death, serious physical injury or injuries requiring hospitalization

Table 10: Percent of CIT Calls by Type

year	mental patient	suicide threat	suicide attempt	suicide
2021	0.8628	0.076	0.0542	0.007
2020	0.886	0.0704	0.0376	0.006
2019	0.8957	0.057	0.0395	0.0078
2018	0.9149	0.0401	0.0287	0.0163
2017	0.946	0	0	0.054

Note: Mental health calls are comprised of calls classified as type "mental patient", "suicide threat", "suicide attempt" or "suicide" by the dispatcher. This table shows the proportion of mental health calls of each of these types by year.

[14] from NOPD policy manual

Table 11: Trimmed Panel

	Dependent variable:	
	force	arrest
All Calls	(1)	(2)
Treated	−0.0004	−0.02
	(0.0003)	(0.002)
Observations	370,259	370,259
R^2	0.016	0.043
Mental Health	(1)	(2)
Treated	0.0018	0.019
	(0.0033)	(0.009)
Observations	11,526	11,526
R^2	0.008	0.012
Note:	*p<0.1; **p<0.05; ***p<0.01	

Treated = 1 if the officer received training and the incident occurred after training was completed. Fixed effects for incident priority, day of week, shift and district are included. The top panel uses the repeated cross-section sample of all valid incidents and the bottom panel uses the sample of only mental health incidents. Only officers who never receive training or receive training 6 months or more after the observation window begins (June 2017-Feb 2020) are included. Standard errors are clustered at the incident level.

Table 12: Do Officer Characteristics Predict Training?

	Dependent variable:
	training received
Age	0.001
	(0.001)
Years of service	-0.014^{***}
	(0.001)
Male	-0.084^{***}
	(0.015)
race: Black	0.064
	(0.065)
race: Hispanic	0.118
	(0.075)
race: Native American	0.220
	(0.220)
race: Two or More	0.872^{***}
	(0.219)
race: Unknown	0.287^{***}
	(0.069)
race: White	0.051
	(0.065)
Observations	4,426
R^2	0.129
F Statistic	72.896^{***} (df = 9; 4416)
Note:	*p<0.1; **p<0.05; ***p<0.01

The dependent variable is a binary indicator for whether officer *i* ever receives training in the window of observation. The sample includes all officers trained in this window and all never-trained officers (so excludes the always-trained officers).

Table 13: Does Training Predict Force and Arrests?

	Dependent variable:	
	force	arrest
	(1)	(2)
cit	0.001***	0.008***
	(0.0001)	(0.001)
Observations	1,345,192	1,345,192
R^2	0.011	0.026
F Statistic (df = 57; 1345134)	260.839***	637.440***
Note:	*p<0.1; **p<0.05; ***p<0.01	

I test whether training predicts use of force or propensity to make an arrest. The sample is limited to incidents involving officers who are never-trained and officers who will be trained prior to the officer receiving train- ing. The dependent variable is an indicator for force use/arrest and the main regressor is an indicator for will-be-trained. Controls for call characteristics are included.

Table 14: Training Cohorts Demographics

Date of Training Completion	Percent Black	Percent Male	Avg Years of Service
Nov 13, 2015	0.54	0.77	16
May 27, 2016	0.47	0.70	15
Aug 26, 2016	0.62	0.47	12
Oct 28, 2016	0.56	0.76	15
Feb 3, 2017	0.38	0.64	12
April 25, 2017	0.02	0.92	5
Oct 13, 2017	0.40	0.77	12
Feb 23, 2018	0.40	0.65	11
June 8, 2018	0.45	0.65	10
Aug 24, 2018	0.27	0.82	9
Nov 30, 2018	0.53	0.70	13
Mar 29, 2019	0.61	0.65	8
June 28, 2019	0.49	0.63	9
Jan 31, 2020	0.61	0.46	7
July 30, 2021	0.55	0.85	6
Dec 10, 2021	0.61	0.89	6

Note:

Table 15: Training Cohorts Use of Force and Arrest Rates

training completion month	force rate	arrest rate
Feb 2017	0.004	0.177
April 2017	0.003	0.241
Oct 2017	0.003	0.200
Feb 2018	0.003	0.181
June 2018	0.003	0.213
Aug 2018	0.003	0.245
Nov 2018	0.003	0.185
March 2019	0.004	0.229
June 2019	0.003	0.213
Jan 2020	0.003	0.207

Note: Columns 2 and 3 are average percentages of incidents officers respond to per month prior to training that result in use of force or arrest.

CIT At its core CIT is de-escalation training for first responders with the underlying goal of diverting those with mental illness from the criminal justice system to the health care system and community services. The Memphis Model of CIT includes 40 hours of training through classroom lectures, interactions with community service providers, mental health care providers, and those with mental illnesses. Lectures cover topics including but not limited to: mental illnesses medications and side effects, co-occurring disorders, developmental disabilities, suicide prevention, personality disorders, and PTSD. De-escalation is taught with scenario-based interaction in five stages: Basic Strategies, Basic Verbal Skills, Stages/Cycle of a Crisis Escalation, Advanced Verbal Skills, and Advanced Strategies: Complex Scenarios. While NOPD states to be following this model, the local NAMI branch is not involved in training, nor is any other local organization, rather training was developed in-house and is done in-house.

Table 16: Mental Health Calls with CIT Response

year	n total	n cit	perc
2017	5,701	3,649	0.640
2018	6,540	4,466	0.683
2019	6,729	3,845	0.571
2020	6,598	2,316	0.351

Note: Col 1 is the total number of mental calls that per given year. Col 2 is the total number of mental calls that has a CIT trained officer respond per given year. Col 3 is quotient of Col 2 and Col 1.

6.1 Static TWFE

I estimate a two-way fixed effects model at the officer-month level. An officer is marked as treated for any incident they are dispatched to after their training date, and all data is pooled into one pre-period and one post-period.

$$Y_{ict} = \alpha_i + \gamma_t + \phi Treated_{ict} + \beta X_c + \epsilon_{ict} \tag{3}$$

Y_{ict} is an indicator for whether officer i used force or made an arrest at incident c in month t. α_i are officer fixed effects and γ_t are time fixed effects. $Treated_{ict}$ is an indicator equal to one if officer i had already received CIT training by month t when responding to incident c. X_c is a vector of call characteristics including district, priority, shift, and day of week. The sample is all calls for service for mental health incidents.

The coefficient of interest is ϕ. Recent developments in the difference-in-differences literature have shown that the estimate of ϕ provides an unbiased estimate of the average treatment effect only if two assumptions hold: parallel trends and constant treatment effects between groups and over time. The first is the classic assumption of difference-in-difference models, while the second emerged in the recent literature.

In this context the parallel trends assumption would require that officers who receive CIT training and officers who never receive training would have maintained common use of force and arrest trends in the absence of treatment. To assess the validity of this assumption I plot the rates of use of force and arrest rates prior to training for officers who receiving CIT training and rates of use of force and arrest rates for officers who never do, presented in Figure 1. If we took the average the never-treated and treated officers would have very similar rates of force use and arrests in the pre-period, but by plotting the pre-trends month-by-month we see they are not parallel. Officers who select into training have higher rates of both force and arrests prior to training than the never-treated officers and the arrest trends cross, potentially violating parallel trends.

Table 17 presents the estimated coefficients on $Treated$ from Equation (2). These estimates imply that CIT training is actually associated with an *increase* in use of force and propensity to arrest. However, to interpret these estimates causally we must assume constant treatment effects across cohorts and parallel trends and this may be violated.

Table 17: Static Two-way Fixed Effects

	Dependent variable:	
	force	arrest
	(1)	(2)
treated	0.005*	0.016*
	(0.002)	(0.007)
Observations	35,917	35,917
R^2	0.02	0.03
F Statistic	0.92	0.93
Note:	*p<0.1; **p<0.05; ***p<0.01	

The coefficient presented is ϕ from Eq.(2): treated = 1 if the officer had already received training at the time of the call. Fixed effects for incident priority, day of week, shift and district are included. The sample is only mental health incidents. Standard errors are clustered at the incident level.

6.2 Dynamic TWFE

Next, I allow for the possibility of heterogeneous treatment effects across cohorts. CIT training had a staggered roll-out (there had been 17 training classes as of 2021) so I estimate an event-study model with two way fixed effects to allow for variation across training cohorts. I estimate the causal impact of CIT on an officer's actions with the following equation:

$$Y_{ict} = \alpha_i + \gamma_t + \sum_k \phi_{kt} D_{ct}^k + \beta X_c + u_{ict} \tag{4}$$

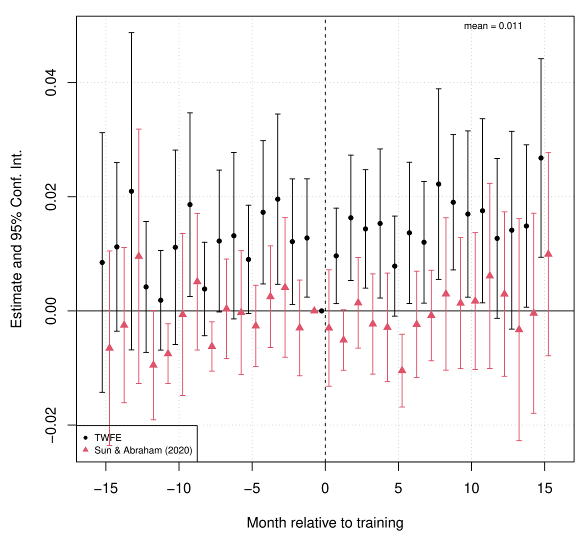
Y_{ict} is an indicator for whether officer i used force or made an arrest at call c in month t. α_i are officer fixed effects, and γ_t are time fixed effects. D_{ct}^k is an indicator equal to 1 if month t is k months after (or before) the month of training, k is negative for the months prior to training and positive for months after training. X_c is a vector of call characteristics including district, priority, shift, and day of week. The sample is all mental health incidents.

Sun & Abraham (2020) provide a decomposition of the population regression coefficients ϕ_{kt}. They show that the coefficients are linear combinations of differences in trends from event period k *and* differences in trends from other relative periods (whether they are included or excluded from the specification). Therefore, group treatment effect estimates may suffer from possible contamination from other event-periods. They propose a method to correct for contamination from other relative periods that weights the average of cohort treatment effects by the shares of cohorts that experience at least k relative treatment periods. Under the joint assumptions of parallel trends and no anticipation the Sun and Abraham interaction-weighted estimator provides an unbiased and consistent estimate of the cohort treatment effect on the treated. CIT training is meant to provide officers with hard and soft skills for managing tense situations and so it is plausible in this scenario to assume no anticipation - just knowing they will be trained will not give them the skills required to de-escalate a volatile situation. Additionally, NOPD has indicated that training is not offered regularly but at various intervals throughout the year, meaning officers cannot predict far in advance when they will be able to sign-up for training and generally are made aware of a new training class only one month (or so) in advance.

As of 2021 there had been 17 cohorts of officers to receive CIT training (1 class in 2015, 3 in 2016, 4 in 2017, 2 in 2018, 2 in 2019, 1 in 2020, and 1 in 2021). In my data I able to observe 9 of these cohorts (2017-2020). To account for the staggered roll-out I use an event-study design. I estimate both a standard two-way fixed effects model as well as the Sun & Abraham (2020) aggregated cohort method.

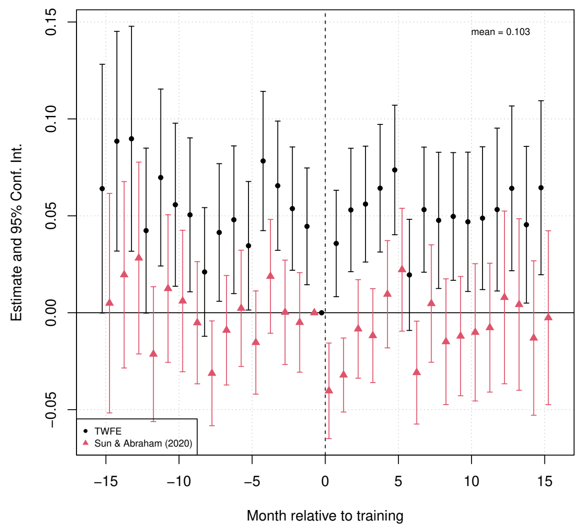
I restrict the sample to 15 pre- and 15 post-periods because the sample is much larger and closer to being balanced in this restricted time frame. Focusing on the Sun & Abraham estimates that account for heterogeneous effects, there are no pretends for force or arrest and the post-trends for both are noisy around zero and so I do not find any impact on use of force or arrests. The event study plots are presented below.

Figure 5: Use of Force



Note: The dependent variable is a binary indicator for use of force. The sample is restricted to only valid mental health incidents. 0 is the month of CIT training completion. Sample includes officers trained during the observation window, never-trained officers and always-trained officers. Standard errors are clustered at the incident level.

Figure 6: Arrest



Note: The dependent variable is a binary indicator for an arrest. The sample is restricted to only valid mental health incidents. 0 is the month of CIT training completion. Sample includes officers trained during the observation window, never-trained officers and always-trained officers. Standard errors are clustered at the incident level.