Mobile Crisis Response Teams Support Better Policing: Evidence from CAHOOTS*

Jonathan Davis
University of Oregon

Samuel Norris
University of British Columbia

Jadon Schmitt
University of Oregon

Yotam Shem-Tov
UCLA and NBER

Chelsea Strickland
Texas A&M

April 2025

Abstract

This paper studies the use of mobile crisis response teams—a non-uniformed pair consisting of a mental health worker and a medic—as a component of emergency response to 911 calls. We provide the first evaluation of the longest-running program in the United States, Crisis Assistance Helping Out on the Streets (CAHOOTS) in Eugene, Oregon, which responds to calls involving mental illness, homelessness, and addiction either instead of or in addition to police officers. We use two complementary research designs to understand the effects and possible scope of these programs. First, we find that a series of program expansions into new areas and times reduced the likelihood that a 911 call resulted in an arrest and increased access to medical services. The arrest reduction likely reflects CAHOOTS' role in de-escalating tense situations and resolving incidents without coercive measures. CAHOOTS is most often dispatched to the same calls as the police, acting as a supplement rather than a substitute. Second, we exploit idiosyncratic variation in CAHOOTS availability in the post-expansion periods to estimate the effect of additional marginal program expansions. We find that they are used mostly for calls that would otherwise go unanswered, suggesting that the program has reached a scale where it can respond to the most urgent calls. We conclude that crisis response teams play an important role as a complement to the police rather than acting only as substitutes.

*Authors can be reached at jdavis5@uoregon.edu, sam.norris@ubc.ca, schmittjadon@gmail.com, shem-tov@econ.ucla.edu, and cstrickland@tamu.edu, respectively. We thank Bocar Ba, Felipe Goncalves, Jens Ludwig, Grant McDermott, Charlie Rafkin, Marit Rehavi, Evan K. Rose, and seminar participants at multiple universities for their helpful comments. Diti Jain and Everett Stamm provided outstanding research assistance. Funding for this project was provided by Arnold Ventures.

Police officers respond to a broad range of 911 calls, spanning violent crimes in progress as well as incidents involving mental health crises, homelessness, addiction, and interpersonal conflicts. However, relying on police as universal first responders may not be optimal, as expanding officers' responsibilities to include mediation and de facto social work could result in strained resources, reduced effectiveness in core law enforcement roles, and unnecessary arrests or violence. An alternative solution that has gained traction in recent years involves integrating non-police first responders into cities' emergency-response portfolios. Typically composed of a non-uniformed mental health worker and a medic dispatched by emergency services, these crisis response teams are trained to address immediate crises arising from substance abuse, mental health issues, and interpersonal conflicts. Advocates argue these situations differ substantially from traditional policing scenarios and thus merit specialized responders (Irwin and Pearl, 2020; Krider et al., 2020).

Despite public demand for alternatives and complements to the police (Ba et al., 2024), there is little evidence on how crisis responders affect the outcomes of 911 calls. Do they substitute the police in responding to calls, or are they dispatched alongside the police as complements? Or do they respond to calls that otherwise would have gone unanswered, expanding the scope of 911 services? Critically, do they successfully de-escalate immediate crises and reduce arrests? While there is an extensive literature on police officers (Anwar and Fang, 2006; Blanes i Vidal and Kirchmaier, 2018; Goncalves and Mello, 2021; Chalfin et al., 2022; Feigenberg and Miller, 2022, 2023; Moreno-Medina et al., 2023; Chalfin and Gonçalves, 2023; Amaral et al., 2023; Ba et al., 2024), the evidence on crisis response teams is sparse (Dee and Pyne, 2022).

This paper investigates the causal impact of sending crisis response teams to respond to 911 calls involving mental illness, homelessness, and addiction—either alongside or instead of the police. Our context is the longest-running such program in the United States, Crisis Assistance Helping Out on the Streets (CAHOOTS) in Eugene, Oregon. We use detailed administrative 911 call data and two complementary research designs to study the effect of the initial expansions and the effect of the marginal call after the program has reached scale, contributing to a comprehensive assessment of the effects and possible scope of crisis response programs.

CAHOOTS originated in Eugene, Oregon in 1989. At the start of our study period in 2014, it was active only in Eugene, running from mid-morning to late at night. Over the following four years, it expanded in five distinct steps to offer 24-hour coverage in Eugene as well as in the neighboring city of Springfield. Our first research design exploits this variation in a difference-in-differences framework, where we instrument for CAHOOTS responses to 911 calls based on the timing of the expansions. This strategy isolates the effect of CAHOOTS responses driven specifically by the introduction to new times and geographies. On average, the expansions increased the probability of a CAHOOTS response by 7.5 percentage points (pp), with a first-stage F-statistic of 178.

We find that CAHOOTS responses induced by the expansions reduced the likelihood that a call resulted in an arrest by 24 pp relative to a control complier mean of 32%. This reduction was driven by arrests authorized under an Oregon law that allows officers to detain individuals who are a risk to either themselves or others because of mental health issues or drug use. This is consistent with descriptive evidence suggesting police responses to mental health crises increase the use of mental health holds (Walker et al., 2019), potentially because they have limited training to diagnose mental health conditions (Cohen and Bagwell, 2023). However, we see no impacts on arrests related to outstanding warrants, suggesting that the 911 center is appropriately targeting crisis response teams towards calls that require their expertise and unique training. Moreover, we find that CAHOOTS improves access to medical services, increasing the likelihood that EMS will arrive, which further supports the interpretation that the averted arrests stemmed from medical rather than criminal causes.

Some advocates of crisis response programs argue that they can serve as a substitute for the police in a wide variety of circumstances, with CAHOOTS itself arguing that they handle 17% of emergency calls and require police backup for only 1% of calls (White Bird Clinic, 2020). Conversely, critics argue that these programs mainly respond to calls that are otherwise outside the purview of the police and would not otherwise receive a response. Our results suggest that neither of these views is correct, and that every call that CAHOOTS responds to as a result of the initial expansions would have otherwise received a police response. However, these responses often arise through CAHOOTS acting as a *complement* to the police and co-responding with them to calls, rather than acting as a substitute.

Our second research design is tailored to identify the impacts of further expanding CAHOOTS availability. To achieve this, we introduce a novel "availability" instrument to assess CAHOOTS' capacity at any given moment. To construct this instrument, we count the number of active CAHOOTS calls when the focal call is received, and define it as high-availability if there is a below-median number of active calls. Through extensive validation exercises, we demonstrate that this measure can be treated as if it were randomly assigned within a specific city, month, day of the week and hour of the day (e.g., Wednesdays 9-10 p.m. in March 2024 in Eugene). The availability instrument strongly predicts whether a call receives a CAHOOTS response with an F-statistic of 526. These instrumented CAHOOTS responses lead to a moderate and statistically insignificant 3.7 pp reduction in the likelihood of an arrest. Furthermore, approximately half of the CAHOOTS responses induced by the availability instrument would have otherwise received no emergency response, whereas nearly all responses induced by the expansion design would have been handled by the police in the program's absence. This suggests that following the expansions—when CAHOOTS responded to roughly 7.5% of 911 calls—the program may have achieved sufficient scale to address the majority of calls that are appropriate for their expertise.

To better understand the differing effects across designs, we develop a simple econometric framework that models the effect of a CAHOOTS response as arising through three possible channels: substitution (CAHOOTS responds instead of the police), complementarity (CAHOOTS responds alongside the police), and service expansion (CAHOOTS responds to calls that otherwise would have gone unanswered). Under natural restrictions on response types, our IV analysis reveals the composition of response types for each design. We show that about half of the CAHOOTS responses induced by the availability instrument would have counterfactually received no other emergency response; in contrast, the 911 center would have dispatched the police to nearly all the CAHOOTS responses induced by the expansions. Since the police are the only emergency responders who are authorized to make arrests, this may explain much of the gap in effects between the expansions and availability designs.

To separately identify the substitution and complementarity effects of CAHOOTS on arrests and quantify their importance, we discuss conditions under which the complier means and corresponding "subLATEs" for each response type can be identified (Kline and Walters, 2016). Under the assumption that the complier means are the same across each of the five expansions—in other words, that differences in the reduced-form effects arise only because of differences in compliance patterns—we can identify the subLATEs. We find that the reductions in arrests arise principally from complementarities between the police and CAHOOTS rather than from a substitution effect, and that this pattern is driven by the high counterfactual arrest risk for calls that are dispatched to both the police and CAHOOTS.

Next, we evaluate the impact of CAHOOTS on access to medical services by linking 911 call data with records of emergency medical services (EMS) responses and assessments. When CAHOOTS is dispatched due to the expansions instrument, the likelihood of EMS arriving on the scene increases by 26 pp, and transports to the emergency room rise by 8.8 pp. Additionally, there is a 4.4 pp increase in the probability of a cardiac arrest diagnosis. In contrast, CAHOOTS responses driven by the availability design show smaller and statistically insignificant effects, mirroring the pattern observed for arrests.

To assess whether CAHOOTS has any adverse effects, we examine its impact on the likelihood of future 911 calls from the same location, which we interpret as a proxy for unresolved problems. CAHOOTS responses reduce the probability of a follow-up call in the following weeks, with larger and statistically significant effects for responses induced by the program's introduction to new areas and hours of the day. These findings suggest that the observed reductions in arrests do not come at the expense of crime prevention or public safety.

Lastly, we combine the estimated benefits and costs of a CAHOOTS dispatch to calculate the marginal value of public funds (MVPF) under each research design. When dispatches are driven by an expansion in coverage, CAHOOTS yields net cost savings—primarily due to lower labor costs relative to police—which implies an infinite MVPF. In contrast, for dispatches triggered by marginal increases in CAHOOTS availability, the MVPF exceeds one as long as the associated social benefits of a response are valued above \$10—which would be satisfied, for example, by valuing an averted arrest at \$270 or more.

This paper contributes to several areas of research. First, there has been much research on police officers and how their behavior and identity impact stops, arrests, and use of force.[1] Another related strand of work examines the impact of police officers on crime and how to optimally allocate officers (Di Tella and Schargrodsky, 2004; Klick and Tabarrok, 2005; Draca et al., 2011; Bazzi et al., 2022; Blattman et al., 2021; Mello, 2019; Premkumar, 2019; Weisburst, 2019; Ba et al., 2021; Chalfin et al., 2022; Mello, 2024; Rivera, 2025a). While increasing the number of police officers has been shown to reduce crime (Evans and Owens, 2007; Chalfin and McCrary, 2018), recent work argues it can also have negative impacts by increasing enforcement of low-level "quality-of-life" offenses, especially for minorities (Chalfin et al., 2022; Hoekstra and Sloan, 2022).

We demonstrate that expanding the range of emergency responders can improve the outcomes of 911 calls by reducing arrests, likely by addressing immediate crises—such as mental health issues—that police are less well-equipped to manage. Most of these improvements are achieved by jointly dispatching police and CAHOOTS teams. This suggests that programs like CAHOOTS can help mitigate the negative effects police may have in certain situations (IACP, 2018; Lum, 2021), benefiting both the individuals involved and the officers responding to the incidents. The de-escalation of immediate crises can have significant benefits given the effect that high-profile incidents of police violence can have on mistrust of the police and levels of criminal behavior (Owens, 2019; Ang et al., 2024; Mikdash and Zaiour, 2024).

Another body of research, primarily consisting of descriptive studies in medical and crimi-

[1] Some recent work includes Miller and Segal (2019); Ba et al. (2021); Hoekstra and Sloan (2022); Rivera (2025b) and Adger et al. (2025), among others.

nology journals, has examined international co-response models, where a police officer and a mental health professional such as a psychiatric nurse or social worker jointly respond to incidents.[2] We contribute to this literature by providing a comprehensive evaluation of the causal effects of CAHOOTS on 911 call outcomes. CAHOOTS is particularly important given the outsized role that it has played as a national template for police reform; cities including Denver, Minneapolis, Olympia, Rochester, and San Francisco (among others) have piloted their own versions. As these programs continue to grow across the country, evidence of their effectiveness and interaction with existing emergency response services is key to their success.

Most related to our study is Dee and Pyne (2022), which examines Denver's Support Team Assisted Response (STAR) program. The program, which was modeled after CAHOOTS, launched in June 2020; a difference-in-differences analysis at the precinct-month level reveals that STAR decreased police reports of the types of offenses targeted by the program (e.g., disorderly conduct, alcohol and drug use, and trespassing) by 34%. We build on their important work in a number of ways. First, our detailed call-level data allow us to disentangle effects on outcomes from changes in reporting induced by more limited police presence. Second, we provide both a reduced-form analysis of the overall effects of CAHOOTS and investigate whether crisis responders should be best understood as a complement or substitute to the police. Third, we study the ways in which crisis response programs increase access to medical treatment, such as by bringing in other services like EMS. Fourth, we use two complementary research designs to evaluate not only the impacts of introducing crisis response teams to a new area but also the scope for the potential gains from additional expansions to program capacity.

2 Background and data

2.1 The crisis response program

The Crisis Assistance Helping Out on the Streets (CAHOOTS) program in Eugene, Oregon was established in 1989 by the White Bird Clinic to address crises involving mental illness, homelessness, addiction, and public assistance. As the acronym slyly suggests, CA-

[2]Studies in the UK have found that co-response teams reduce the number of individuals detained in custody under a mental health hold (Heslin et al., 2016; Jenkins et al., 2017; Keown et al., 2016) and increase hospital admissions (Jenkins et al., 2017). Research in Canada comparing police-only responses to co-responses shows that co-responses reduce police use of force and hospital transports and increase referrals to community services (Blais et al., 2022). In the United States, Helfgott et al. (2016) and Morabito et al. (2018) discuss the implementation of co-response programs in Seattle and Boston, respectively. Shapiro et al. (2015) and Marcus and Stergiopoulos (2022) provide systematic reviews.

HOOTS operates as a partnership with police and other emergency services in the area. Calls to 911 or the police non-emergency line are routed through the Lane County emergency communications center, which handles dispatch for police, fire, and ambulance, as well as CAHOOTS.

A CAHOOTS team consists of two people—a medic and a mental health crisis worker—in a Sprinter-style van loaded with emergency supplies. Medics may be a nurse, paramedic, or (most often) an EMT.[3] Crisis workers are not necessarily licensed professionals, but have at least two years of experience in a non-traditional mental health setting, such as working with homeless outreach programs or crisis phone lines. Instead of police training, they receive specialized instruction on topics such as de-escalation, harm reduction, behavioral health, and substance abuse. Following classroom instruction, new team members complete 500 hours of field training, in which the trainee observes (and is supervised by) two experienced team members. Upon successful completion of field training, the trainee begins working unsupervised in their own two-person team. Team members are unarmed and do not wear a uniform. The team's goal during each call is to address callers' immediate needs, rather than to provide long-term solutions. This might include, for example, transporting callers experiencing a mental health crisis to either a health care facility or a friend's house, as the situation warrants. However, CAHOOTS also connects individuals to available programs such as drug rehabilitation support.

2.2 The dispatching decision process

At any given time, the Lane County emergency communications center is staffed by several calltakers who are the public's first point of contact after they phone 911 or the police non-emergency line. The calltaker's job is to speak with the caller to understand the situation and record relevant information. The computer-aided dispatch (CAD) system provides information on the address of the call and previous calls from the same number, as well as information about the caller in arrest and court records. Usually within seconds of talking to the caller, the calltaker records the nature of the call from a pre-populated list of options; the most common call natures by response type are shown in Table 1. Calltakers then assign a numerical priority to the call. While calltakers also record extensive notes in the CAD system, we do not have access to this information.

[3]Emergency Medical Technicians (EMTs) provide injury assessments, emergency medical care, and transportation to medical facilities. Compared to paramedics and nurses, EMTs are more limited in the type of care they can provide. For example, EMTs can deliver oxygen to patients, administer CPR, and provide Narcan for overdoses, but are not allowed to perform intubations or administer intravenous fluids.

Sitting in the same room as the calltakers are the dispatchers. The dispatcher's job is to communicate with emergency responders and assign specific units to calls. As the calltaker talks to the caller and adds information into the CAD system, the dispatcher decides which unit—the specific police, fire, ambulance, or CAHOOTS vehicle—to assign to the call, and how to prioritize it given the currently outstanding calls. The lowest priority calls are often not responded to—for example, a welfare check where the caller phones back to say it is no longer required. After the call has received a response or it has been deemed that a response is no longer necessary, the call is closed.

2.3 Data sources and sample construction

Our analysis relies on computer-aided dispatch (CAD) data from the Eugene and Springfield police departments, which we link to arrest records for each city. Data for Eugene and Springfield includes all calls between January 1, 2014 and December 14, 2021, including both emergency and non-emergency calls. We also have data on medical- and fire-related 911 incidents that we use to construct some call outcomes.

We observe detailed information on each call including the date and time, nature (e.g., suicidal person, theft, robbery), priority, and location. Response information includes primary and secondary responders, time of each unit's dispatch and arrival, time of closing, and whether the call resulted in an arrest.[4] We merge call data with census tract information. For each census tract, we observe a large number of characteristics including the size of the population, gender and racial makeup, median home values and rents, employment and unemployment rates, as well as per capita and household income.

For calls with an arrest, we also observe a summary of the charges used to justify the arrest. We use this information to separately look at arrests for involuntary mental health holds, detox, and warrants. Mental health holds require that an individual be identified as having a mental disorder and as being a threat to themselves or others (Gagnon et al., 2022); detoxification holds work similarly for intoxicated individuals.

Our primary objective is to examine the substitutability and complementarity between CAHOOTS and police responses and how these dynamics influence call outcomes. Although our data include both 911 and non-emergency calls, our analysis of response patterns (i.e., whether CAHOOTS or the police are dispatched) and outcomes focuses on 911 calls. We exclude non-emergency calls because they are less likely to result in a police dispatch.

[4] We also observe reported use of force, but it is an extremely rare outcome and we do not have sufficient statistical power to detect effects.

2.4 Summary statistics

Table 1 reports information on 911 calls in Eugene and Springfield from January 1, 2014 to December 14, 2021. CAHOOTS and the police are responsible for very different types of emergencies. Panel A reveals that CAHOOTS teams disproportionately respond to calls for a wellness check and suicidal persons, but are not assigned to go alone to potentially violent situations such as robberies, fights, armed persons, or drunk drivers. The police respond to a large number of traffic incidents (Panel B), as well as disorderly subjects, suspicious persons, and potential trespassing and burglaries. However, Panel C reveals that CAHOOTS also responds to more potentially serious incidents alongside the police; 6.4% of disorderly subjects calls receive a response from both the police and CAHOOTS.

These differences in types of calls reflect the differing specialties and training of CAHOOTS and the police. CAHOOTS is able to provide a range of on-scene services above and beyond those typically offered by the police, including first-aid and non-emergency medical care, suicide assessment and intervention, conflict resolution and mediation, and counseling for people grappling with grief and loss, substance use, homelessness, and other crises.

Table 2 reveals that there are also differences in call outcomes. Panel A shows that since CAHOOTS is not authorized to make arrests, there are no arrests on CAHOOTS-only calls. In contrast, 14.4% of police-only calls and 13.3% of CAHOOTS-and-police calls result in an arrest, with much higher rates of non-criminal and detoxification holds among the CAHOOTS-and-police calls. Panel B shows that 18.7% of CAHOOTS-and-police calls receive any fire/EMS response compared to 10.9% of police-only calls and 4.3% of CAHOOTS-only calls.

Finally, Panel C of Table 2 examines call responses. Regarding the time to response, the first unit takes longer to arrive on CAHOOTS-only calls than calls to which the police are also sent. The average time to response is just under 14 minutes for a CAHOOTS-only call and about 6 minutes for a police-only call. This reflects the fact that CAHOOTS calls are often less urgent and that there are fewer CAHOOTS vans available. Once on the scene, CAHOOTS stays for about 26 minutes for a CAHOOTS-only call and for about 42 minutes when they accompany the police. Police units cumulatively spend 90 minutes on the scene when responding alone and 152 minutes when responding with CAHOOTS, although the per-unit response durations are lower. In calls where both CAHOOTS and the police are dispatched, the police almost always arrive first (87%) and then CAHOOTS joins later, potentially because the police requested their assistance.

3 Effects of expansions in CAHOOTS services

3.1 CAHOOTS expansions research design

Our first research design seeks to quantify the impact of introducing CAHOOTS services on call outcomes. Specifically, at five distinct times during our study period, CAHOOTS vans suddenly became available for additional hours in either Eugene or Springfield.[5] Table A1 contains a list of these policy changes, which we use to estimate the effect of a CAHOOTS expansion in a difference-in-differences framework.

Figure 1 plots the number of 911 and non-emergency calls in each month for the relevant geography and hours of the day. For ease of interpretation, we normalize the number of calls by the number of hours in the design; for example, the first policy change in Eugene affects three hours (7-10 a.m.), so we divide the total number of calls each month by three. The changes in the number of calls are sharp and sudden, with the number of calls per month jumping at the time of expansion by 30-50 per hour-month in Eugene and 10-15 in Springfield.[6] Each of the changes are permanent, with the number of calls slightly increasing over the subsequent years. Focusing only on 911 calls, Figure A1 shows similar patterns. The number of CAHOOTS 911 calls per hour-month for each policy change jumps sharply and discontinuously at the timing of the expansions.

These graphs inform our first empirical strategy, which leverages the expansions as instrumental variables for CAHOOTS dispatch. Specifically, we employ a simple difference-in-differences specification, which for a single expansion is structured as follows:

$$\begin{aligned} Y_i &= \beta \overbrace{C_i}^{\substack{\text{CAHOOTS}\\ \text{response}}} + \psi^1_{t(i)} + \gamma^1_{c(i)} + u_i \\ C_i &= \alpha \underbrace{Z_i}_{\substack{\text{CAHOOTS}\\ \text{expansion}}} + \underbrace{\psi^0_{t(i)} + \gamma^0_{c(i)}}_{\substack{\text{Calendar month}\\ \text{\& city FEs}}} + \varepsilon_i \end{aligned} \tag{1}$$

where Y_i is an outcome such as a recorded arrest as a result of call i, C_i is an indicator for a CAHOOTS response, and Z_i is an indicator for the call coming in a time and geography following a CAHOOTS expansion (i.e., "Post $\times$ Treatment" in the standard difference-in-

[5] There were several policy changes in Eugene that added or removed a van, but without ever moving the number of available vans to zero. These policy changes had a much less dramatic effect on the likelihood that a call was answered by CAHOOTS, so we focus on the extensive margin changes.

[6] Before each expansion, the number of calls that were dispatched to CAHOOTS was very small but non-zero, and was the result of either vans working past the end of their shift to address additional calls, or dispatchers waiting for CAHOOTS vans to begin their shifts so they could assign a call to them.

differences notation). We control for calendar month (e.g., May 2018) effects $\psi_{t(i)}$ and city effects $\gamma_{c(i)}$.

We observe five expansions and we therefore use a stacked design that combines all policy changes each as an instrumental variable for CAHOOTS. Since some of the designs overlap—for example, Table A1 reveals that the first Eugene expansion and the second Springfield expansion both included 9-10 a.m.—we additionally control for the other policy changes. Each design j consists of the hours of the day affected by the policy change, whether or not they are in the city where the policy change occurred. Specifically, we estimate the following 2SLS specification:

$$Y_{ij} = \beta C_{ij} + \psi^1_{t(i)j} + \gamma^1_{c(i)j} + \sum_{j' \neq j} \pi_{jj'} Z_{ij'} + u_{ij} \tag{2}$$
$$C_{ij} = \alpha_j Z_{ij} + \psi^0_{t(i)j} + \gamma^0_{c(i)j} + \sum_{j' \neq j} \alpha_{jj'} Z_{ij'} + \varepsilon_{ij}$$

where Y_{ij} is an outcome on call i in design j and Z_{ij} is an indicator for the call coming in a time and city following the j^{th} expansion. $t(i)$ indexes months and $c(i)$ the two cities in our sample, so we control for design-month effects $\psi_{t(i)j}$ and design-city effects $\gamma_{c(i)j}$, analogous to time and location fixed effects in a traditional difference-in-differences specification. Some observations appear twice in this regression because of the design overlap, so we cluster standard errors at the call level. The stacked specification in Equation 2 combines all five expansions to a single over-identified 2SLS specification and can be interpreted as a weighted average of the IV estimates from each expansion separately.[7]

3.2 Balance checks

We begin by showing that there are no changes to the type or composition of 911 calls that coincide with the timing of the expansions (e.g., as a response to changing need for services). Figure 2 reports the coefficients from a regression of indicators for the calltaker descriptions of the type of call on the expansions instrument. We find no effects on these call natures, nor on predicted measures of a police response, arrest, or an involuntary hold. Moreover, in robustness analysis below, we show that controlling for these predicted call natures does not impact our estimated effects.

[7] The weights are proportional to the number of observations in the design multiplied by the residual variance in the instrument-predicted treatment. In our reported reduced form regressions, we use a single post-expansion indicator and variance-weight across designs to match the implicit design weights in the IV specfication. This procedure ensures that our reduced form regressions are consistent with our main 2SLS specification and implies that if they are re-scaled by the first stage, they provide the same CAHOOTS effect estimate as our main 2SLS specification.

Figure 2 also reports the coefficients from an OLS regression of the call natures and predicted outcomes on a CAHOOTS response. In contrast to the IV approach, there are large and precisely estimated correlations between a CAHOOTS response and the type of call. This highlights that simple comparisons between calls with and without a CAHOOTS response may be misleading due to potential confounding factors.

3.3 The effect of CAHOOTS on call responses and arrests

Panel A of Table 3 reports reduced form estimates of the effect of an expansion on call outcomes. Consistent with Figure 1, the expansions caused a dramatic change in the response composition, with the likelihood of a CAHOOTS response jumping by 7.5 pp (F=178).[8] Simultaneously, column (2) shows that the likelihood of a police response decreased by 3 pp, suggesting that in about 40% of the instrument-induced responses CAHOOTS acted as a replacement for the police. Column (3) reveals that there was nearly no effect on whether there was any response, suggesting that the remainder of the CAHOOTS responses—about 60%—were alongside the police. The remaining columns in Panel A show that there were declines in the number of arrests.

To recast the reduced form estimates in terms of a CAHOOTS response, Panel B turns to estimating 2SLS regressions of the effect of a CAHOOTS response on call outcomes. Consistent with Panel A, column (2) of Panel B reveals that a CAHOOTS response reduces the likelihood of a police response by 39.6 pp relative to a control complier mean (CCM) of 96.6%, indicating that the expansion-induced CAHOOTS responses crowd out a considerable number of police responses. However, as shown in column (3), there is nearly no effect on whether there is any response, with only 3.4% of marginal CAHOOTS responses counterfactually receiving no response (the other 96.6% would have received a police response).

The remainder of Panel B estimates the effect on call outcomes. A CAHOOTS response dramatically reduces the likelihood of arrest, by 24 pp relative to a CCM of 32%. This reduction is almost entirely explained by involuntary holds (15 pp) and detoxification holds (5 pp), suggesting that CAHOOTS reduces arrests mostly by offering voluntary alternatives to non-criminal detention. We also find that CAHOOTS has no effect on warrant arrests, consistent with the 911 center appropriately dispatching them to relatively low-risk calls. The CCM of arrest is more than three times the rate in non-CAHOOTS calls (see Panel C), indicating that expansion-induced CAHOOTS responses are directed toward calls where

[8]Before the expansions, CAHOOTS responded to about 0.5% of calls. These responses occurred either when the caller phoned early in the morning and was placed on hold by the dispatcher until CAHOOTS started work for the day, or when CAHOOTS worked past their regular hours.

there is substantial scope to improve outcomes and where, in the absence of CAHOOTS, coercive measures are more likely to be used.

Our analysis also reveals evidence of substantial omitted variable bias consistent with the imbalances documented in Figure 2, suggesting that simple comparisons of 911 calls with and without CAHOOTS can be misleading. For instance, OLS estimates in Panel C suggest that CAHOOTS increases involuntary hold arrests—an effect that is opposite in direction to the 2SLS estimates. Moreover, OLS underestimates the impact of CAHOOTS on arrest, showing only a 3.8 pp reduction in arrests, compared to 2SLS estimates that show a substantively larger 24 pp reduction in arrests.

3.4 Robustness

We conduct a number of analyses to validate and check the robustness of the results in Table 3. First, we show our estimated effects are robust to the inclusion of additional controls. While our information on each call is limited, we observe the exact address, allowing us to map calls to their 2010 Census tract and measure a large number of tract characteristics from the ACS, including the size of the population, gender and racial makeup, median home values and rents, employment and unemployment rates, as well as per capita and household income. Table A2 shows the results, with Panel A replicating our preferred specification. Panel B controls for these tract characteristics, and Panel C additionally adds the call nature as coded by the calltaker. Panel D instead includes tract-city-design fixed effects. Across each of these robustness exercises, we find remarkably similar results; for example, our estimate of the effect of a CAHOOTS response on arrests ranges only from -22.6 pp to -24.1 pp.

Second, to further validate the expansion instrument, we test whether placebo event times affect arrest rates and find no significant effects. Figure A3 displays the distribution of estimated effects for 500 randomly chosen event times. These effects are estimated using the same specification as in Equation 2, focusing on the reduced-form relationship. The placebo event time effects are small, centered around zero, and do not overlap with the estimated effect from the observed data. These results are consistent with the estimates in Figure 2 showing that the the CAHOOTS expansions are not correlated with changes in the nature of the calls or with tract characteristics.

Third, Table A3 presents reduced-form estimates corresponding to Equation 1 for each expansion separately. Reassuringly, all expansions increase the likelihood of a CAHOOTS response and decrease the likelihood of an arrest. Although the policy changes differ substantially in how much they raise the probability of a CAHOOTS response, the effects on arrests are similar for the Eugene and Springfield expansions. This consistency aligns with the over-identification J-test, which does not reject the null hypothesis that all expansions recover a common estimate of CAHOOTS' effects on arrests ($p = 0.835$; see Table 3).

Fourth, one might worry that there were reductions in police capacity at the time of the the CAHOOTS expansions, which might directly affect calls. Figure A2 shows that there were no appreciable changes in the number of units on duty at the time of the expansions. More directly, Table A7 controls for a measure of police availability (see Section 4) and finds almost identical results.

Finally, we assess robustness to alternative ways of constructing counterfactual outcomes by estimating the reduced-form effect using both the traditional two-way fixed effects designs analogous to Equation 2 and the imputation approach of Borusyak et al. (2024). Since this approach uses only ever-treated city-hour-of-day cells, the sample is much smaller. Still, the results are very similar (see Table A4), further bolstering the robustness of our results.

4 Effects of marginal CAHOOTS responses

4.1 Availability research design

Using the expansions of CAHOOTS to new times of the day and geographies provides an estimate of the effect of introducing CAHOOTS as an additional option relative to a counterfactual where the police are the only available emergency service.[9] In this section, we study the likely effect of further additions to CAHOOTS capacity by estimating the effect of marginal CAHOOTS responses induced by quasi-exogeneous CAHOOTS availability *after* the initial expansions.

Our approach is to instrument for a CAHOOTS response with an indicator for there being a below-median number of active 911 calls to which CAHOOTS units are already assigned when the focal call is received. Specifically, for each focal call, we count the number of prior 911 calls that are still open separately for CAHOOTS- and police-dispatched calls, which—after residualizing out city-month-day of week-hour of day fixed effects to account for variation in available units and normalizing to have a mean of zero and a standard deviation of one—we call busyness (W_i). We then create a binary instrument $Z_i \equiv 1[W_i \leq \text{median}(W_i)]$ that indicates above-median CAHOOTS availability.[10] Because of the fixed effects, these

[9] This is not strictly true since there is a non-zero share of calls that are responded to by CAHOOTS before the expansions. However, this occurs very rarely—in only 0.5% of calls—and only when callers wait for CAHOOTS to become available or the CAHOOTS teams work overtime.

[10] Using a binary instrument makes the comparison to the expansions design results from Section 3 more

comparisons focus, for instance, on 911 calls received in Springfield between 10 and 11 PM on Wednesdays in March 2021, comparing times with a high versus low level of CAHOOTS availability. We use a 2SLS specification, similar to Equation 1, in which a CAHOOTS response is instrumented by our measure of CAHOOTS availability:

$$\begin{aligned} Y_i &= \beta C_i + \gamma^1 W_i^p + \eta^1_{h(i)} + \varepsilon_i \\ C_i &= \alpha \underbrace{Z_i}_{\substack{\text{High CAHOOTS} \\ \text{availability}}} + \gamma^0 \underbrace{W_i^p}_{\substack{\text{Police} \\ \text{busyness}}} + \eta^0_{h(i)} + \varsigma_i \end{aligned} \tag{3}$$

where Y_i is an outcome such as a recorded arrest during call i, Z_i is the availability instrument, W_i^p is the number of open police calls (i.e., police busyness), and $\eta_{h(i)}$ is fixed effects at the city-month-day of week-hour of day level. Our design therefore compares calls received from the same area and within a very tight temporal window, but that were received at times of differing CAHOOTS availability.

4.2 First stage and balance checks

Figure 3 visualizes the first stage effect of the availability instrument on the likelihood of a CAHOOTS response using the continuous measure of CAHOOTS availability. The relationship is negative and monotonic; an additional standard deviation of busyness (approximately 3.5 additional calls) increases the likelihood of a CAHOOTS response by 2.1 pp (F-statistic = 453). A natural concern is that CAHOOTS availability is related to call characteristics even conditional on our rich set of fixed effects. To assuage this concern we conduct a battery of placebo tests where we regress observable call characteristics on CAHOOTS availability using the reduced-form analog to the IV specification.

The estimates in Figure 2 show that calls that come at times with more availability are not associated with differences in call natures. Each row reports the coefficient on call availability (blue solid square markers) or CAHOOTS assignment (red hollow circles) in a regression with the outcome given in the row header. CAHOOTS availability is uncorrelated with call natures and predicted response or outcomes. However, these characteristics are strongly predictive of actual assignment; as indicated by the hollow red circle markers, almost all of the characteristics are related to CAHOOTS assignment even conditional on the fixed effects. The predictive power of the call natures for actual assignment suggests that our placebo exercises should have some power to detect violations of exogeneity. This is particularly clear when the outcome is the likelihood of arrest as predicted using census tract

straightforward and is consistent with the conceptual econometric framework that we discuss in Section 5.1.

characteristics and call natures. A CAHOOTS response is cross-sectionally associated with a 2 pp lower predicted arrest rate (mostly reflecting the difference between calls that do and do not receive any sort of response), but the call coming at a time of high availability has a negligible and statistically insignificant effect on the predicted arrest rate of 0.125 pp.

4.3 Effect of marginal CAHOOTS responses on call outcomes

Panel A of Table 4 reports reduced form estimates of the effect of high CAHOOTS availability on call outcomes. Consistent with Figure 3, the availability instrument leads to a sharp increase of 3 pp in the likelihood of a CAHOOTS response with an F-statistic of 526. Unlike the expansions instrument, marginal increases in CAHOOTS responses largely reflect answering calls that otherwise would have gone unanswered; higher CAHOOTS availability increases the likelihood of responding to a call by 1.6 pp.

Panel B of Table 4 turns to estimating 2SLS regressions of the effect of a CAHOOTS response on call outcomes. Column (2) reveals that a CAHOOTS response reduces the likelihood of a police response by 17.3 pp relative to a CCM of 45.9%, indicating that the availability-induced CAHOOTS responses crowd out a considerable amount of police response. However, as shown in column (3), there is a large effect on whether there is any response, with 54.1% of marginal CAHOOTS dispatches counterfactually receiving no response.

The remainder of Panel B estimates the effect on call outcomes. A CAHOOTS response decreases the likelihood of arrest by a statistically insignificant 3.7 pp. This coefficient is statistically different than the estimate derived using the expansion instrument, suggesting potential heterogeneity in effects. Consistent with the muted effects on overall arrests, we also find small effects of a marginal CAHOOTS response on involuntary and detoxification holds, as well as warrant arrests. At a high level, we take these results to mean that there might be smaller benefits from additional increases to CAHOOTS capacity than there were from making it available at new times of the day.[11]

Finally, Panel C of Table 4 examines the difference between the OLS and 2SLS estimates, and finds clear differences. For example, OLS estimates show a large increase in arrests involving mental hold, contrary to the 2SLS estimates which show a null effect. We interpret this difference to mean that the OLS estimates are likely driven by omitted variable bias given the significant imbalances in call types documented in Figure 2.

[11] Table A5 assesses the robustness of the estimated effects of CAHOOTS to the inclusion of additional controls. As shown in the table, the results remain stable when controlling for census tract characteristics, call nature codes, and tract-by-city-by-design fixed effects, further reinforcing the credibility of the estimates.

5 Estimating complementarity and substitution

5.1 An econometric framework for interpreting effects

The effect of a CAHOOTS response triggered by higher-than-usual availability is a moderate but statistically insignificant reduction in arrests. This contrasts with the CAHOOTS responses prompted by the expansions which, as discussed in Section 3, reduce arrests by nearly six times as much. A possible explanation for these differing effects is variation in the counterfactual responses. As discussed in the previous sections, about half of the CAHOOTS responses induced by the availability design would counterfactually receive no response; the corresponding number in the expansions design is almost exactly zero. However, there may be other differences across designs, such as whether they induce CAHOOTS to respond alongside the police. In this section we present a simple econometric framework to clarify these dynamics.

We split the responses to a 911 call into four possible options: no response (n), only CAHOOTS is dispatched (c), only the police are dispatched (p), and both CAHOOTS and the police are dispatched (b). Let $D_i \in \mathcal{D} \equiv \{n, c, p, b\}$ denote the response to call i, and let $Y_i(d)$ denote the potential outcome of interest under treatment d. To tractably describe the different counterfactuals we will focus on the case of a binary instrument $Z_i \in \mathcal{Z} \equiv \{0, 1\}$ that increases the likelihood of a CAHOOTS response. For a given instrument, there are therefore $|\mathcal{D}|^{|\mathcal{Z}|} = 16$ potential *response types*, or types defined by the set of counterfactual treatments for each instrument value (Heckman and Pinto, 2018).

Guided by our institutional setting, we make two restrictions on the allowable response types. First, we assume that the CAHOOTS instrument increases the likelihood of a response that involves CAHOOTS for each individual, analogously to the usual monotonicity assumption (Imbens and Angrist, 1994). Second, we assume that the instrument does not affect the likelihood of a police response except in the case of substitution. For example, Z_i can move calls from p to c, but not from p to n.[12] These restrictions can be formally stated as:

Assumption 1 (IA monotonicity)**.**

$$\mathbb{1}[D_i(1) \in \{c, b\}] \geq \mathbb{1}[D_i(0) \in \{c, b\}]$$

[12]We can relax this assumption to be conditional on a measure of the busyness of the police. We use a measure of police busyness similar to our CAHOOTS availability instrument as an additional control in robustness analysis below.

Assumption 2 (No cross effects except substitution)**.**

$$\begin{aligned}
\mathbb{1}[D_i(1) \in \{b\}] &\geq \mathbb{1}[D_i(0) \in \{b\}] \\
\mathbb{1}[D_i(1) \in \{n, c\}] &\geq \mathbb{1}[D_i(0) \in \{n, c\}] \\
\mathbb{1}[D_i(1) \in \{c, p, b\}] &\geq \mathbb{1}[D_i(0) \in \{c, p, b\}]
\end{aligned}$$

The allowable response types are summarized in Table 5. Under these assumptions, the instrument changes treatment for three groups: calls that are moved from p to b, calls that are moved from p to c, and calls that are moved from n to c. Each of these response types therefore faces a different change in their treatment: for the $p{\to}b$ compliers CAHOOTS is a *complement* to the police, for the $p{\to}c$ compliers CAHOOTS is a *substitute* to the police, and for the $n{\to}c$ compliers there is a *service increase* since their calls would otherwise not receive a response.[13] The Wald estimator of the effect of a CAHOOTS response can therefore be decomposed as:

$$\frac{E[Y|Z{=}1] - E[Y|Z{=}0]}{E[\underbrace{1[D \in \{c,b\}]}_{\substack{\text{CAHOOTS}\\\text{response}}}|Z{=}1] - E[1[D \in \{c,b\}]|Z{=}0]} = \underbrace{\frac{\pi_{pb}}{\pi_{nc} + \pi_{pc} + \pi_{pb}}}_{\equiv\, s_{pb}} \underbrace{\overbrace{E[Y(b) - Y(p)|p \to b]}^{\equiv \Delta_{pb}}}_{\substack{\text{Adding CAHOOTS to}\\\text{a police response}}} \tag{4}$$

$$+ \underbrace{\frac{\pi_{pc}}{\pi_{nc} + \pi_{pc} + \pi_{pb}}}_{\equiv\, s_{pc}} \underbrace{\overbrace{E[Y(c) - Y(p)|p \to c]}^{\equiv \Delta_{pc}}}_{\substack{\text{Substituting from}\\\text{police to CAHOOTS}}} + \underbrace{\frac{\pi_{nc}}{\pi_{nc} + \pi_{pc} + \pi_{pb}}}_{\equiv\, s_{nc}} \underbrace{\overbrace{E[Y(c) - Y(n)|n \to c]}^{\equiv \Delta_{nc}}}_{\substack{\text{CAHOOTS response}\\\text{instead of no response}}}$$

where we denote compliers of type $D_i(0){=}d$, $D_i(1){=}d'$ as $d{\to}d'$ and where the shares of the complier groups, $\pi_{dd'} \equiv P[D(0){=}d, D(1){=}d']$, can be identified from changes in observed treatments:

$$\begin{aligned}
\pi_{pb} &= E\big[\mathbb{1}[D_i{=}b] \mid Z_i{=}1\big] - E\big[\mathbb{1}[D_i{=}b] \mid Z_i{=}0\big], \\
\pi_{pc} &= E\big[\mathbb{1}[D_i{\in}\{n,c\}] \mid Z_i{=}1\big] - E\big[\mathbb{1}[D_i{\in}\{n,c\}] \mid Z_i{=}0\big], \\
\pi_{nc} &= E\big[\mathbb{1}[D_i{\in}\{c,p,b\}] \mid Z_i{=}1\big] - E\big[\mathbb{1}[D_i{\in}\{c,p,b\}] \mid Z_i{=}0\big].
\end{aligned} \tag{5}$$

Equation 4 highlights that differences in 2SLS estimates using different instruments can be the result of differences in the treatment effects Δ_{pc}, Δ_{pb}, and Δ_{nc}, or due to differences in

[13] Assumption 2 is particularly intuitive in the case of the expansions design. This is because there are no effects on whether there is any response (suggesting there are no $n{\to}p$ or $n{\to}b$ compliers), and because the instruments cause an extensive margin shift in CAHOOTS (so there can be no compliers with $D(0){=}c$).

the composition of complier types s_{pc}, s_{pb}, and s_{nc}. As a first step towards characterizing possible differences in treatment between the availability and expansion designs, Figure 4 reports the shares of compliers from each response type, $s_{dd'} = \pi_{dd'}/(\pi_{nc} + \pi_{pc} + \pi_{pb})$. We find sharp differences between the designs. Expansion-induced CAHOOTS responses are complements ($p \to b$) in 57% of calls and substitutes ($p \to c$) for 39.6% of calls; they expand services ($n \to c$) in only 3.4% of calls. In contrast, for the availability-induced CAHOOTS responses, 53.8% of calls would not have been responded to, with sharp declines in the number of both complement and substitute compliers.[14]

We view these results as suggesting that there are a relatively limited number of 911 calls where CAHOOTS can operate as a substitute for the police, and after the initial expansions—which boosted the share of CAHOOTS calls from approximately 0 to 7.5%—most of these calls have been exhausted. Instead, most of the availability-induced CAHOOTS responses are to calls that otherwise would not have received a response. Further expansions to CAHOOTS capacity—by adding additional vans during the day, for example—would likely match this pattern.

A rescaled version of Equation 4 provides additional insight into possible treatment effect heterogeneity across the expansions and availability designs. In our setting, where the $n \to c$ effect on arrest must be zero, we can recover the LATE of a weighted average of Δ_{pc} and Δ_{pb} under the same identifying assumptions using the following Wald estimator:

$$\frac{E[Y|Z=1] - E[Y|Z=0]}{E[1[D \neq p]|Z=1] - E[1[D \neq p]|Z=0]} = \underbrace{\frac{s_{pb}}{s_{pc} + s_{pb}} \cdot \Delta_{pb} + \frac{s_{pc}}{s_{pc} + s_{pb}} \cdot \Delta_{pc}}_{\text{Rescaled effect}}. \tag{6}$$

Panel A of Table A6 reports 2SLS estimates of the rescaled effect from Equation 6. Relative to the non-rescaled estimates in Table 3, the rescaled estimates of the effect in the expansions design are very similar because of the negligible number of $n \to c$ compliers. However, in the availability design, the rescaled estimates in Table A6 are roughly twice as large in magnitude as the non-rescaled estimates in Table 4 because of the high level of $n \to c$ compliers. Nonetheless, the rescaled effects remain quite different for the availability and expansion designs (-8.6 pp and -24.1 pp, respectively). This cannot be explained by differences in the weights on Δ_{pc} and Δ_{pb}, which as Panel B reports are similar across designs. We conclude

[14] One testable implication of Assumption 2 is that each of the estimated π's should be positive; Figure 4 shows that this condition is satisfied. Our assumptions also implies that these shares should be positive in each subsample. In Figure A4, we estimate them separately for the periods of day affected by each of the expansions and the figure clearly shows that the estimated πs are non-negative in both the expansions (Panel A) and availability (Panel B) designs.

that there is likely some heterogeneity in treatment effects across the expansions and availability designs. In the next section, we explore this further by providing direct evidence on the subLATEs and the potential outcomes underlying them.

5.2 Cross-design differences in complier outcomes

Given our model, we show in Appendix A1 that information about outcomes for the complier types is contained in the following two moments:

$$\frac{E\big[Y1[D{=}b \mid Z{=}1]\big] - E\big[Y1[D{=}b \mid Z{=}0]\big]}{E\big[1[D{=}b \mid Z{=}1]\big] - E\big[1[D{=}b \mid Z{=}0]\big]} = E[Y(b) \mid p{\to}b] \tag{7}$$

$$\frac{E\big[Y1[D{=}p \mid Z{=}1]\big] - E\big[Y1[D{=}p \mid Z{=}0]\big]}{E\big[1[D{=}p \mid Z{=}1]\big] - E\big[1[D{=}p \mid Z{=}0]\big]} = \frac{s_{pb}}{s_{pb}+s_{pc}}E[Y(p) \mid p{\to}b] + \frac{s_{pb}}{s_{pb}+s_{pc}}E[Y(p) \mid p{\to}c]$$

where s are the known complier shares.[15]

The first moment reveals that the data directly identify the average likelihood of arrest when both the police and CAHOOTS respond for the $p{\to}b$ compliers. In contrast, the second moment implies that the data only contain information on a weighted average of $Y(p)$ for $p{\to}b$ and $p{\to}c$ compliers, rather than directly identifying each response group-specific $E[Y(p) \mid g]$. These average potential outcomes, however, are informative for understanding the differences in results across designs.

Table 6 contains estimates of the moments in Equation 7 for both the expansions and availability designs. Panel A reveals that $E[Y(b) \mid p{\to}b]$ is nearly identical across designs; if both the police and CAHOOTS respond as a result of either instrument, the likelihood of arrest is about 14%. In contrast, the average value of $Y(p)$ for $p{\to}b$ and $p{\to}c$ compliers varies wildly across the designs, from 0.315 in the expansions design to 0.169 in the availability design. This difference cannot be explained by the weights on the complier means; Panel B of Table 6 reveals that they are almost identical across designs. Instead, we view this as evidence that the marginal call changed dramatically from the time of the expansions—when CAHOOTS responded to nearly no calls—to the post period, when CAHOOTS had reached scale. The CAHOOTS responses induced by the expansion had a very high risk of arrest in the counterfactual where the police responded. After the expansions, however, the remaining marginal calls—which are captured by the availability instrument—were less pressing, and so had a relatively low risk of arrest.

[15] The likelihood of arrest is zero for $D \in \{n, c\}$, which implies that $E[Y(c) \mid p{\to}c] = E[Y(c) \mid n{\to}c] = E[Y(n) \mid n{\to}c] = 0$. In Appendix A1, we show that this means that the identifying variation in the instruments can be simplified from eight moments reflecting the Wald estimator of the effect of the relevant treatment on $Y1[D{=}d]$ or $Y1[D{\neq}d]$ to only the two in Equation 7.

5.3 Disentangling substitution and complementarity

Since the previous analysis does not separately identify $E[Y(p) \mid p{\to}b]$ and $E[Y(p) \mid p{\to}c]$, it does not provide any direct evidence on the relative magnitude of the complementarity effect Δ_{pb} and the substitution effect Δ_{pc}. One way to view this is that the treatment effects are simply underidentified: a single binary instrument provides only two moments to identify three parameters.

To overcome this difficulty, we note that one could generate the moments in Equation 7 for each of the five expansions. Then, under the assumption that the complier means are the same for each expansion, one can directly estimate these complier means using a minimum distance procedure to fit the system of ten linear equations in three unknowns. Table 6 shows that we cannot reject the null hypothesis that $E[Y(b) \mid p{\to}b]$ is the same across the five expansions, consistent with constant compliers means.

Table 7 contains the results. Panel A reveals that the average arrest rates with a police-only response differ dramatically across the $p{\to}b$ and $p{\to}c$ groups: nearly 50% of the former group would be arrested by the police while less than 10% of the latter group would be.[16] We view this difference as reflecting the varying severity of calls across these response groups; the calls that the dispatchers add CAHOOTS to after they become available are much higher risk than the calls where they substitute for the police.

These differences in complier means correspond to very different complementarity and substitution effects. Panel B of Table 7 reports these effects; substituting the police with CAHOOTS reduces the arrest rate by 8.9 pp (SE=18.9 pp) while complementing the police with CAHOOTS reduces the arrest rate by 36.3 pp (SE=16.4 pp). We conclude that much of the arrest-reducing benefit of CAHOOTS comes from working together with the police, rather than as an alternative.[17,18] These findings suggest that crisis response teams can function as effective complements to the police—an important dimension that has received little attention in scholarly and policy discourse.

[16] Note that the 0.133 estimate for $E[Y(b) \mid p{\to}b]$ is slightly different than the 0.136 estimate in Table 6; this is because the latter is estimated via 2SLS rather than variance-weighted minimum distance.

[17] Recall that Table 2 shows that police arrive first 88 percent of the time when both police and CAHOOTS respond.

[18] Panel C of Table 7 reports the results of additional over-identification tests, both overall and for the $E[Y(p)]$ moment. For both tests we find p-values above 0.99, consistent with our assumption of constant complier means across expansions.

5.4 Discussion

A final question is whether the subLATEs estimated using the expansions are consistent with the effects from the availability design. To analyze this, in Panel C we predict the rescaled availability design effect using the subLATEs from the expansions design and the availability response type weights. These estimates—which mechanically eliminate differences in response type composition—are substantially larger in magnitude than the empirical estimates of the availability design in Panel A—26.3 pp vs. 8.6 pp. We take this as additional evidence that the arrest-reducing benefits of the availability-induced CAHOOTS responses are likely smaller than in the initial expansion, even among calls facing the same change in treatment. This suggests that at its current scale, dispatchers are already able to allocate CAHOOTS to the calls where they are most likely to improve outcomes. Further increases in capacity are likely to have smaller effects.

6 Effects on access to medical services

Emergency responders are able to request additional emergency services if needed when they are responding to a call. Our analysis so far has focused on emergency services provided by the police and CAHOOTS. In this section, we measure the impact of a CAHOOTS response on whether a medical unit also responds to the call and on indicators of health outcomes and access to medical services.

Panel A of Table 8 shows estimates of the impact of a CAHOOTS response on access to medical services using our expansion instruments. Sending CAHOOTS increases the probability that a fire or EMS unit also responds to a call by 26 pp, relative to a control complier mean close to zero. These additional fire/EMS responses lead to a statistically significant 4.4 pp increase in the probability that a cardiac arrest is identified and a statistically significant 8.8 pp increase in the probability that an individual is transported to the ER, relative to control complier means that are indistinguishable from zero. We view this as evidence that one of the benefits of crisis response teams might be more effective deployment of other emergency services and expansion of access to needed medical services.

Panel B of Table 8 shows that the descriptive correlation between a CAHOOTS response and these medical outcomes, controlling for the same controls in our expansions research design, are quite different than the causal effects implied by our 2SLS estimates. The descriptive correlations indicate that a CAHOOTS response is associated with a 1 pp *reduction* in fire/EMS responses, increases in suicidal persons, overdoses, and a reduction in ER transports. Such naive OLS comparisons even show that CAHOOTS is associated with a higher likelihood that an individual is dead on arrival.

Next, we examine the impact of a marginal CAHOOTS response using our availability research design. As discussed above, CAHOOTS is more likely to provide an additional service in these cases rather than acting as a complement or substitute to a police response. Panel C of Table 8 presents estimates of how these marginal responses affect health and medical outcomes. We find that they increase the probability of identifying a suicidal person by 4.7 pp, or 82% relative to the CCM of 5.7%. However, they do not significantly affect the likelihood of an additional fire or EMS response, ER transport, or identifying an overdose or a cardiac arrest.

Mirroring the results in Panel B of Table 8, Panel D shows that the association between a CAHOOTS response and these health outcomes, controlling for the same set of controls used in our availability design, looks quite different from the causal effect of marginal responses. We take this as another reminder of the risk of confounding variables in descriptive comparisons.

7 Effects on future calls from the same address

Our analysis up until now has revealed that dispatching CAHOOTS can have a profound effect on the call itself, with a CAHOOTS response substantially reducing the likelihood of an arrest, particularly in the expansion research design. We measure the need for future 911 services using an indicator for a future call from the same address within a given number of days, excluding the first 24 hours, to avoid capturing anything related to the focal call itself. We interpret this measure as a potential indicator of an unresolved problem, as well as future crime.

The red line in Panel A of Figure 5 reports the results using the expansions IV design. CAHOOTS appears to *reduce* the need for future emergency services in the medium term; the point estimates are all negative and statistically significant until the 33$^{\text{rd}}$ day after the focal call, after which they get smaller. The longer-run point estimates indicate a reduction in future calls; however, the effects are not statistically significant. The short-run effects are relatively large in magnitude, although not precise, with a CAHOOTS response reducing the likelihood of a future call within two weeks by about 25 pp relative to a control complier mean of about 55% (Panel B).[19]

[19] To probe the robustness of our measure of future calls, Figure A5, Panel A measures future calls using anything received within 50 meters of the focal call, while Panel B shows the estimated impact on future calls within 50 meters of the focal call, excluding the exact location. In both cases, we see declines in the

Next, the blue line in Figure 5 reports the impacts of CAHOOTS responses induced by our availability instrument on future calls from the same address. The point estimates are negative for the first 19 days after the focal call was received; however they are mostly statistically insignificant and—consistent with the smaller effect on arrests—never more than half of the magnitude of the effects in the expansions design. In the longer term, the point estimates are close to zero and are statistically insignificant. We take this as evidence that even on the margin, CAHOOTS responses may reduce the need for future emergency services in the short-run and provide no evidence of adverse long-run effects of CAHOOTS responses.

The analysis of future calls from the same location helps address a possible concern that in the long term CAHOOTS is counterproductive, particularly if the police would have appropriately deterred those individuals likely to require a future emergency response. Our estimates suggests this is unlikely.

8 Welfare and cost trade-offs

CAHOOTS offers significant cost savings when it replaces police. However, since many CAHOOTS responses either complement the police or expand services, the net costs of a crisis responder program are unclear. In this section, we introduce a simple marginal value of public funds (MVPF; Finkelstein and Hendren, 2020) framework to quantify these costs and weigh them against the potential benefits. We also quantify some of the broader benefits of CAHOOTS on emergency response operations.

Our MVPF analysis focuses on weighing the possible social benefits of reduced arrests against the labor cost of a CAHOOTS response. We calibrate the cost of a response as two responders $\times$ time of response $\times$ wage per hour. Since the average time on call (excluding time traveling to the call) is almost exactly half an hour for both CAHOOTS and each police unit, this amounts to simply the hourly wage for each type of responder. We peg these wages at \$46.34 for the police[20] and \$18 for CAHOOTS (Eugene Register Guard, 2024). These measures almost certainly meaningfully understate the cost difference between CAHOOTS and the police, given the pensions and other benefits for which police officers are eligible. They also likely understate the additional labor cost of processing an arrest, given that the average 911

likelihood of a future call. Finally, Panel C excludes generic addresses such as intersections and entire blocks, as well as locations with more than 300 911 calls; we also see declines in the likelihood of a future call from these addresses where displacement may be less likely.

[20] The sixth step of the pay scale for police officers; see https://www.eugene-or.gov/DocumentCenter/View/2387/City-of-EugeneSalary-Schedule.

call involving an arrest takes about 3 hours.

We combine the cost and benefits of a CAHOOTS response using the following MVPF formula

$$MVPF = \frac{-V_{arrest}\beta^{arrest}}{s_{nc}\kappa_c + s_{pb}\kappa_c + s_{pc}(\kappa_c - \kappa_p)}.$$

where s denotes the complier shares, κ_d denotes the cost of response d, V denotes the social value of an averted arrest, and β the effect of a CAHOOTS response on the likelihood of arrest.

The denominator is an estimate of the impact of a CAHOOTS dispatch on the cost of a 911 response: the cost of a CAHOOTS response for $n{\to}c$ and $p{\to}b$ compliers and the difference in the cost between CAHOOTS and police for $p{\to}c$ compliers. Weighting by the complier shares in Figure 4 suggests that a CAHOOTS response induced by an expansion in CAHOOTS services saves about 35 cents per call. In contrast, a marginal expansion in CAHOOTS services costs $9.98.

The numerator is an estimate of the marginal willingness to pay for the benefits of a CAHOOTS response. We conservatively assume all of the social benefits of CAHOOTS are driven by their impact on arrests. Because CAHOOTS responses induced by expansions in CAHOOTS services are cost-saving, the MVPF will be infinite so long as the value of the benefits is positive. The MVPF of responses induced from marginal expansions will be greater than one if the social value of prevented arrests is at least $269.82. While evidence on the appropriate social value of an averted arrest is sparse, Cohen and Piquero (2009) estimate the criminal justice costs alone for an arrest for "other offenses" is about $750.[21,22] We conclude that given the low costs of even the marginal expansions, the MVPF of CAHOOTS is likely greater than one throughout the range of service levels we observe in our data.

The MVPF analysis leaves out several additional hard-to-value benefits of CAHOOTS. For example, we do not attempt to incorporate the value derived from better connecting individuals to appropriate medical services. Perhaps more importantly, we do not account for possible improvements in emergency response time arising from the CAHOOTS program. Figure A6 demonstrates that CAHOOTS significantly reduces response times under both research designs. Although the benefits are difficult to value, faster response times reduce the amount of time that other callers have to wait for service and may even improve clearance

[21] $500 in 2007 inflated to 2024 dollars.

[22] This is also roughly comparable to the labor cost of making an arrest, since police units spend just over three hours on calls that result in an arrest.

rates (Blanes i Vidal and Kirchmaier, 2018).[23]

Lastly, the MVPF does not account for distributional benefits. CAHOOTS predominantly responds to calls in poorer neighborhoods, making it a progressive public service. Figure A7 displays the descriptive correlation between the share of 911 calls CAHOOTS responds to in a census tract and average rent prices (Panel A) and the share of non-White residents (Panel B). This figure indicates that, in addition to its benefits in reducing arrests and improving access to medical services, CAHOOTS primarily benefits low-income and marginalized populations.

9 Conclusion

We provide new evidence on the effectiveness of CAHOOTS, a crisis response program in Eugene, Oregon that has served as a model for cities across the country grappling with emergencies involving homelessness, mental health crises, and addiction. Using rich 911 call data, we find that using CAHOOTS teams as both substitutes and complements for the police decreases the likelihood of arrest on the focal call and increases access to medical services. It also expands access to other city services and appears to be an effective medium-term solution, decreasing the likelihood of a subsequent call from the same address over the next several weeks.

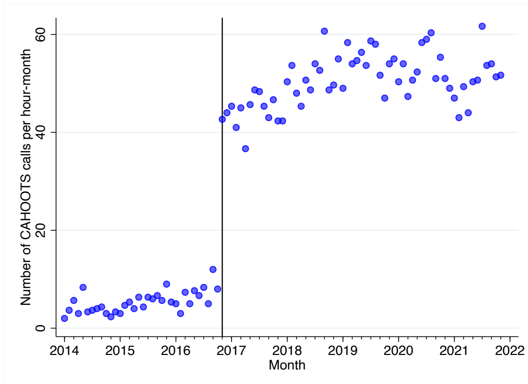
However, CAHOOTS' ability to serve as a substitute for the police is limited. We find that in the initial expansions that made CAHOOTS suddenly available in new times of day or geographic areas, about 40% of CAHOOTS responses were substitutes for the police (in the other 60%, CAHOOTS responded alongside the police). However, after the initial expansions, when CAHOOTS responds to about 7.5% of calls, most of the calls where CAHOOTS can serve as either a substitute or complement have been exhausted. In particular, marginal responses prompted by idiosyncratic availability of CAHOOTS units are mostly to calls that otherwise would have received no response. Future research should determine the extent to which these patterns are caused by statutory restrictions on the types of calls that CAHOOTS can respond to, or whether there are more general limits to the crisis responder model.

[23] The reduction in response times raises the question of whether our instruments impact arrests also by reducing police response times. To probe the robustness of our estimates, we examine specifications that control for a measure of police availability analogous to our CAHOOTS availability instrument. Table A7 shows that our effects remain robust after including this additional control, which, if anything, strengthens the estimated impacts of CAHOOTS.

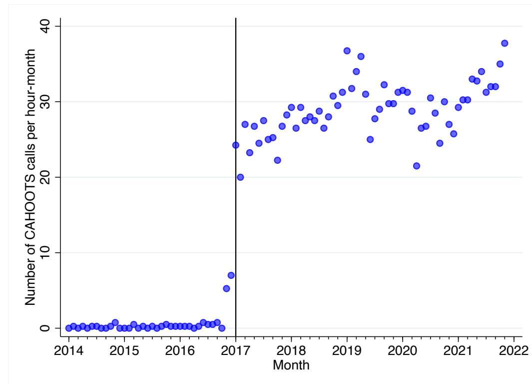
Figures

Figure 1: CAHOOTS expansions in Eugene and Springfield

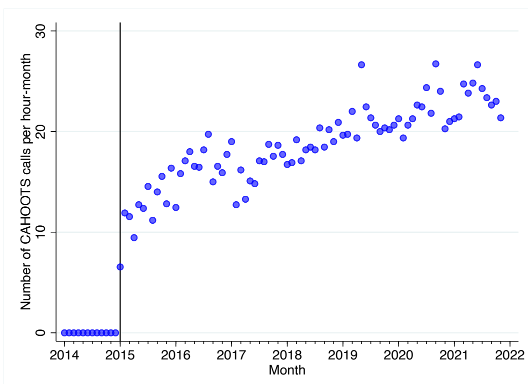
(a) Eugene 7-10 a.m.



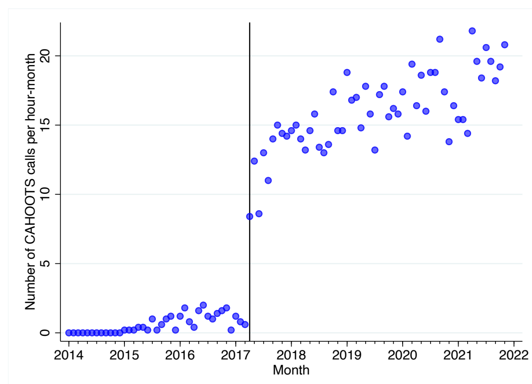
(b) Eugene 3-7 a.m.



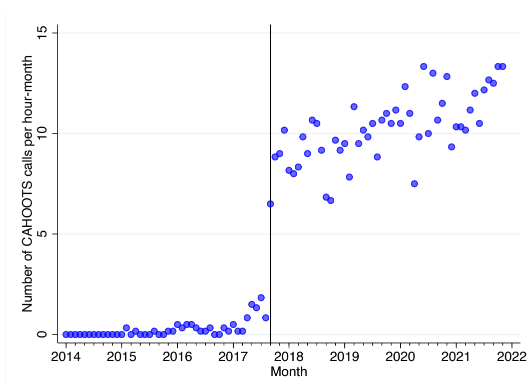
(c) Springfield 12 p.m. to 11 p.m.



(d) Springfield 9-11 a.m. and 11 p.m.-2 a.m.

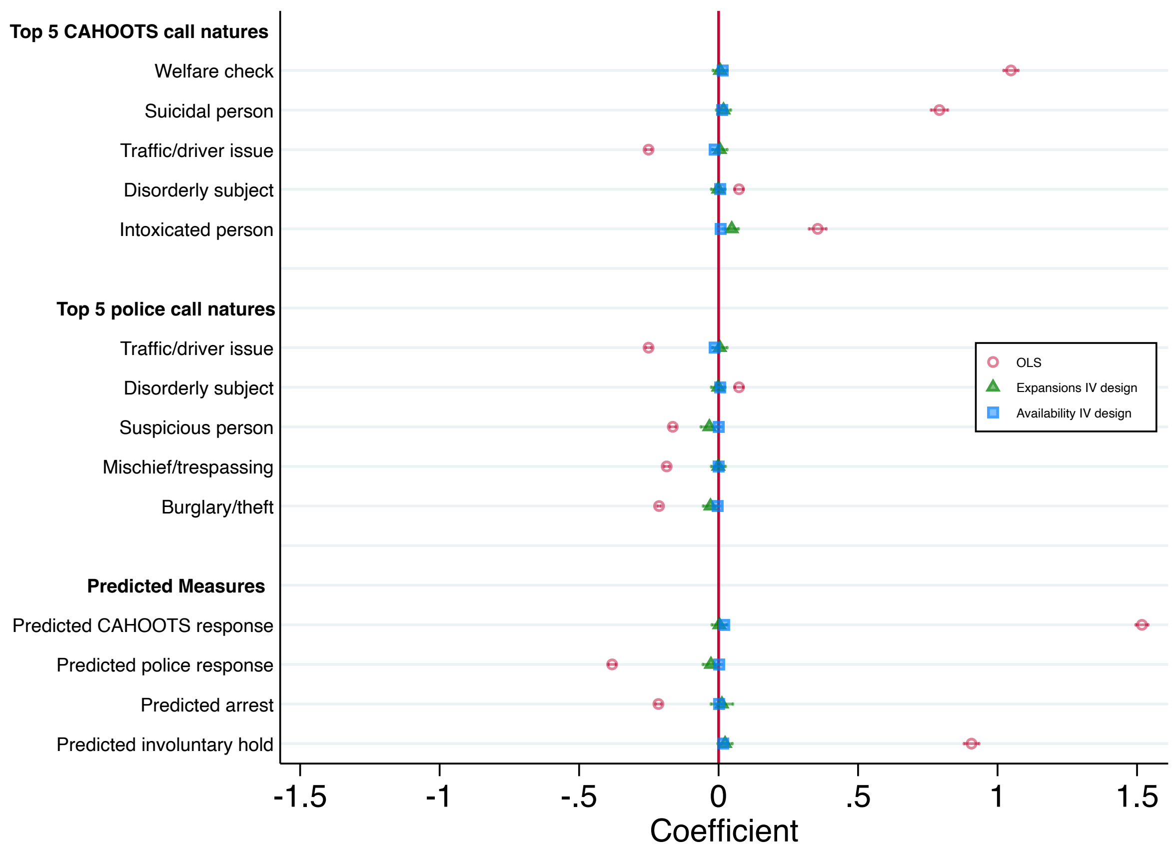


(e) Springfield 3-10 a.m.



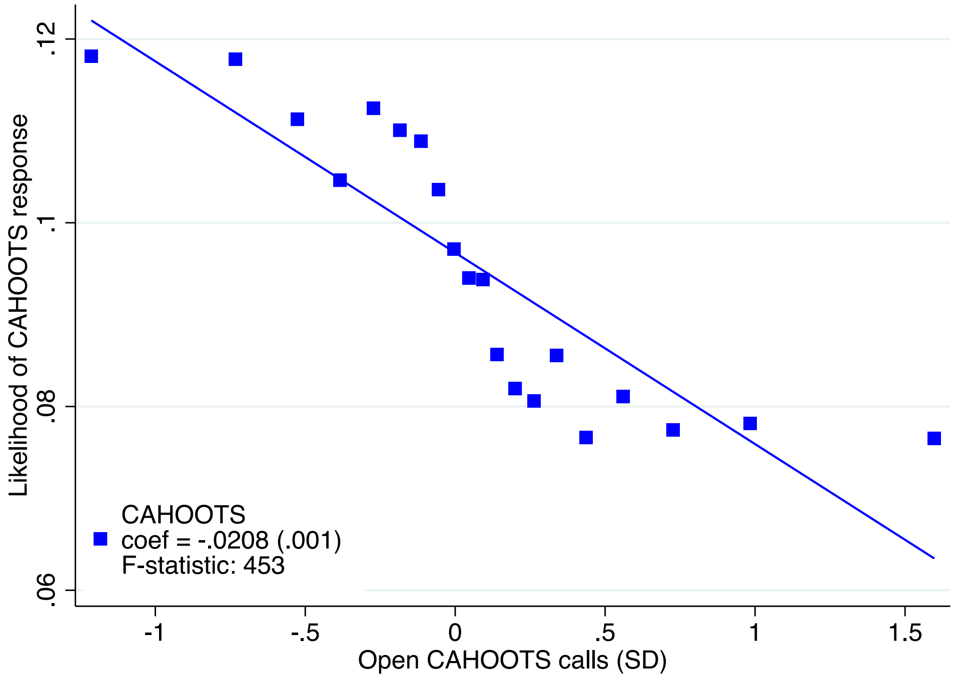
Notes: This figure shows the changes in the number of calls to which CAHOOTS is dispatched as services are expanded (Eugene) or introduced (Springfield). Panels A and B report expansions in Eugene and Panels C, D, and E expansions in Springfield. Each point in the figures reports the number of calls (911 and non-911) to which CAHOOTS was dispatched, normalized to be measured in terms of calls per hour-month. Before each expansion, the number of calls that were dispatched to CAHOOTS was very small but non-zero, and were the result of either vans working past the end of their shift to address additional calls, or dispatchers waiting for CAHOOTS vans to begin their shifts so they could assign a call to them.

Figure 2: Balance checks of CAHOOTS expansions and availability instrument



Notes: The figure reports estimates of the association between CAHOOTS response and call natures as well as predicted measures based on natures and the geographic characteristics of the call location (red hollow circle markers). It also includes the association between CAHOOTS expansions and call characteristics (green solid triangular markers), and between the binary CAHOOTS availability instrument and call characteristics (blue solid square markers). All outcome variables have been standardized to have a mean of zero and a standard devision of one to be on a comparable scale for the figure.

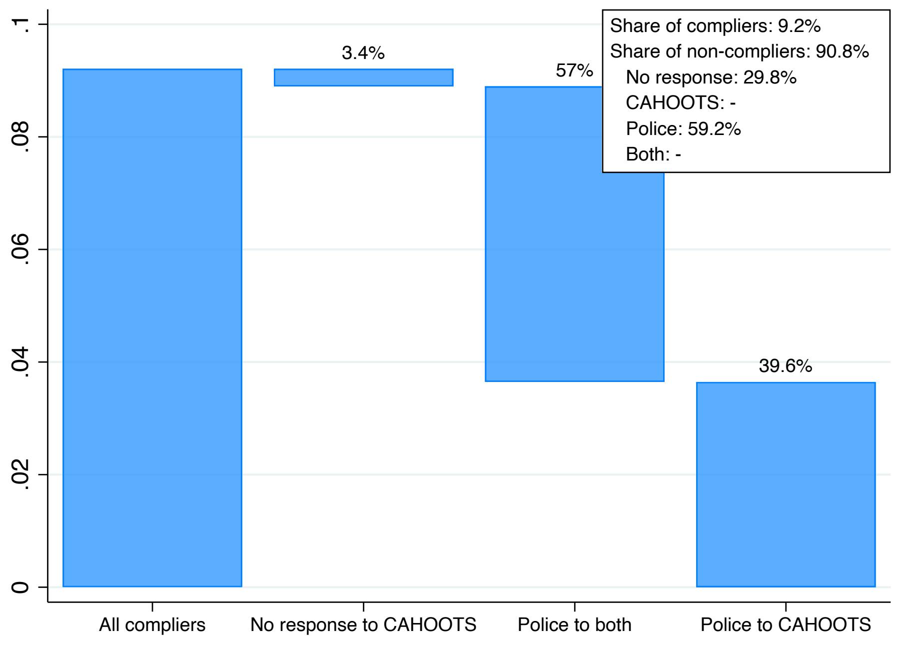
Figure 3: Busyness and the likelihood of a CAHOOTS response



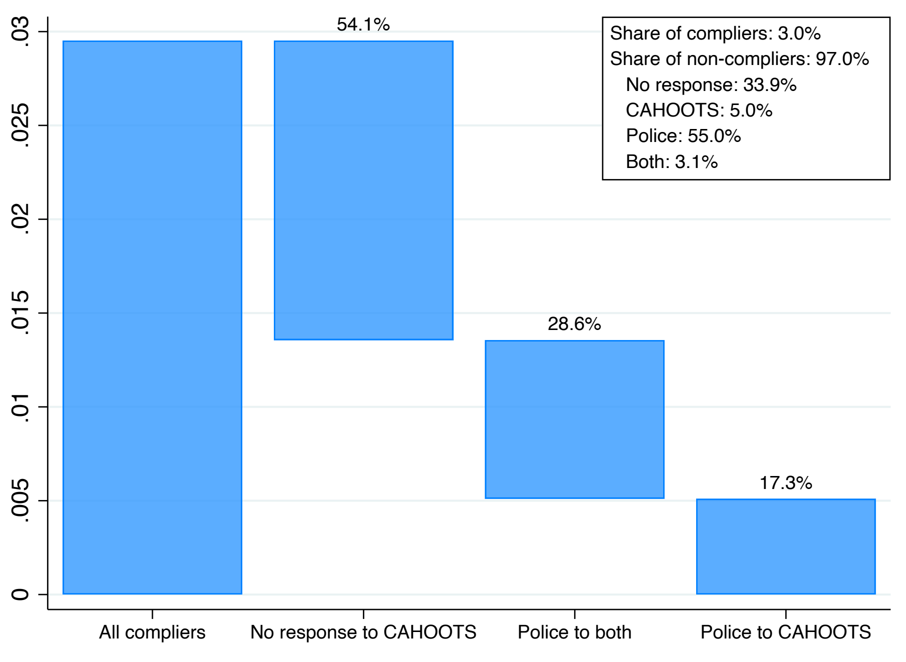
Notes: The figure shows the first-stage relationship between the standardized busyness measure for CAHOOTS and the likelihood of getting a CAHOOTS response.

Figure 4: Shares of response types in expansions and availability designs

(a) Expansions design



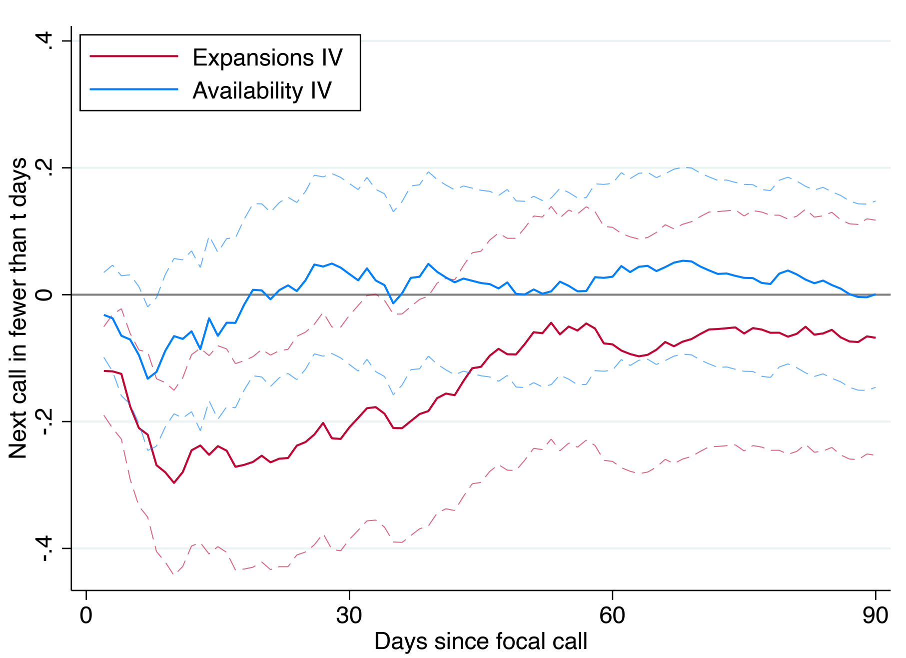
(b) Availability design



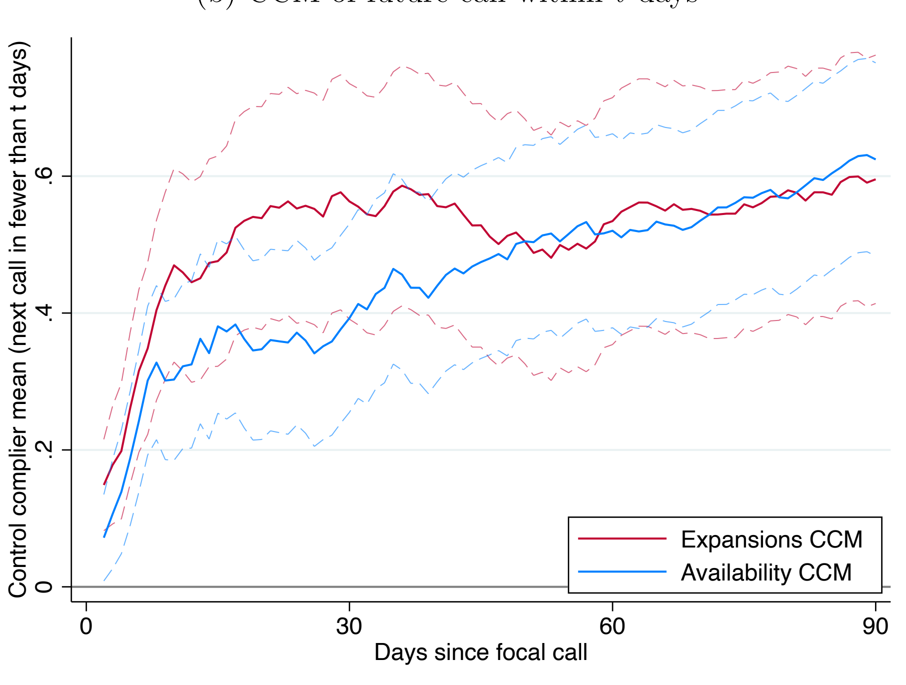
Notes: This figure reports the shares of compliers from each response type using the expansions (left panel) and availability (right panel) designs. The complier share is estimated using a precision weighted average of the first-stage relationships between having any CAHOOTS response and the instruments. The complier type shares are estimated using the instrumented impact of any CAHOOTS response on an indicator for having any response, for having both police and CAHOOTS respond, or for not having a police response, respectively.

Figure 5: Effect of CAHOOTS response on future call from same address

(a) Effect on future call within t days



(b) CCM of future call within t days



Notes: This figure shows the effect of CAHOOTS response on future 911 calls from the same address within a period of time up to 90 days. In Panel A, each coefficient comes from a 2SLS regression of an indicator for a future 911 call within t days on a CAHOOTS response, where CAHOOTS response is instrumented using either indicators for each CAHOOTS expansion or the availability instrument indicator. Panel B contains control complier means. 95% confidence intervals clustered at the individual level are shown in dotted lines.

Tables

Table 1: Call characteristics by response

	(1)	(2)	(3)	(4)	(5)
		Response composition			
	Count	No response	CAHOOTS only	Police only	Both
All calls	270,728	0.334	0.054	0.579	0.033
Panel A: Top 5 most frequent CAHOOTS natures					
Welfare check	16,206	0.217	0.354	0.369	0.060
Suicidal person	6,559	0.131	0.245	0.370	0.254
Traffic/driver issue	31,779	0.490	0.023	0.478	0.009
Disorderly subject	20,070	0.209	0.044	0.683	0.064
Intoxicated person	1,399	0.204	0.415	0.313	0.067
Panel B: Top 5 most frequent police natures					
Traffic/driver issue	31,779	0.490	0.023	0.478	0.009
Disorderly subject	20,070	0.209	0.044	0.683	0.064
Suspicious person	17,097	0.243	0.014	0.721	0.023
Mischief/trespassing	13,251	0.324	0.007	0.649	0.020
Burglary/theft	11,269	0.369	0.001	0.622	0.008
Panel C: Top 5 most frequent natures with both police and CAHOOTS responding					
Dispute	43,042	0.103	0.002	0.864	0.032
Disorderly subject	20,070	0.209	0.044	0.683	0.064
Welfare check	16,206	0.217	0.354	0.369	0.060
Suspicious person	17,097	0.243	0.014	0.721	0.023
Overdose	4,018	0.423	0.031	0.471	0.075
Panel D: Top 5 least frequent natures for CAHOOTS responses					
Robbery	868	0.094	0.001	0.879	0.025
Fight	1,688	0.094	0.000	0.892	0.014
Drunk driver	9,735	0.624	0.000	0.374	0.002
Armed subject	1,409	0.103	0.001	0.869	0.026
Warrant	1,451	0.001	0.000	0.971	0.028

Notes: This table shows the likelihood of each type of response, overall and by call nature, for 911 calls in Eugene and Springfield. The data run from January 1, 2014 to December 14, 2021. Column (1) displays the count of calls. Columns (2)-(5) display the proportion of calls receiving no response, a CAHOOTS only response, a police only response, or a joint CAHOOTS and police response. Panels A, B, and C list call information for the five most frequent call natures receiving a CAHOOTS response, a police response, or a joint CAHOOTS and police response, respectively. Panel D lists call information for the five call natures least likely to receive a CAHOOTS response.

Table 2: Call outcomes by response

	(1) Count	(2) No response	(3) CAHOOTS only	(4) Police only	(5) Both
Panel A: Call escalation outcomes					
Arrest	270,728	0.000	0.000	0.144	0.133
Detox	270,728	0.000	0.000	0.002	0.006
Non-criminal hold	270,728	0.000	0.000	0.010	0.041
Warrant	270,728	0.000	0.000	0.027	0.014
Other arrest	270,728	0.000	0.000	0.105	0.072
Panel B: Health outcomes					
Any Fire/EMS response	270,728	0.105	0.043	0.109	0.187
Transport to ER	270,728	0.034	0.009	0.032	0.035
Dead on arrival	270,728	0.000	0.000	0.001	0.002
Suicidal	270,728	0.007	0.098	0.015	0.170
Overdose	270,728	0.018	0.009	0.011	0.032
Cardiac event	270,728	0.013	0.001	0.010	0.030
Panel C: Call responses					
Time to response	270,728	0.000	13.846	6.308	6.464
Total police time on call	270,728	0.000	0.000	90.170	151.688
Total CAHOOTS time on call	270,728	0.000	25.641	0.000	42.138
Police arrived first	270,728	0.000	0.000	1.000	0.879
CAHOOTS arrived first	270,728	0.000	1.000	0.000	0.120

Notes: This table shows average call outcomes by response type for 911 calls in Eugene and Springfield. Data runs from January 1, 2014 to December 14, 2021. Column (1) displays the count of calls. Columns (2)-(5) display the proportion of calls receiving no response, a CAHOOTS only response, a police only response, or a joint CAHOOTS and police response. Panels A and B display arrest and health outcomes. Panel C displays information on call responses.

Table 3: Effects of CAHOOTS on call response and arrest using expansions design

	Call responses			Call outcomes			
	CAHOOTS (1)	Police (2)	Any response (3)	Arrest (4)	Invol. hold (5)	Detox (6)	Warrant (7)
Panel A: Reduced form effect of expansion							
Post expansion	0.075***	-0.030***	0.003	-0.018***	-0.011***	-0.004***	0.002
	(0.003)	(0.006)	(0.006)	(0.005)	(0.002)	(0.001)	(0.002)
Untreated mean	0.005	0.719	0.721	0.118	0.018	0.005	0.017
Panel B: IV estimates of effect of CAHOOTS response							
CAHOOTS		-0.396***	0.034	-0.240***	-0.151***	-0.051***	0.023
		(0.084)	(0.082)	(0.061)	(0.023)	(0.013)	(0.026)
Control complier mean		0.966	0.966	0.317	0.172	0.054	-0.012
First-stage F-stat		177.76	177.76	177.76	177.76	177.76	177.76
J test of overidentification							
χ^2 statistic		11.087	15.092	1.450	17.174	1.380	5.445
p-value		0.026	0.005	0.835	0.002	0.848	0.245
Hausman test							
χ^2 statistic		1.978	16.692	11.443	59.188	19.196	4.908
p-value		0.160	0.000	0.001	0.000	0.000	0.027
Panel C: OLS estimates of effect of CAHOOTS response							
CAHOOTS		-0.245***	0.375***	-0.038***	0.010***	0.001***	-0.012***
		(0.004)	(0.001)	(0.002)	(0.001)	(0.000)	(0.001)
Untreated mean		0.635	0.635	0.091	0.006	0.002	0.017
Observations		293,845	293,845	293,845	293,845	293,845	293,845

Notes: This table presents estimates of the impact of a CAHOOTS response on call responses and call outcomes using our expansions design. Panel A reports reduced form estimates of the effect of an expansion on these call outcomes. Panel B shows 2SLS estimates of the effect of a CAHOOTS response on call outcomes using indicators for each expansion as instruments. Panel C shows the association between a CAHOOTS response and these outcomes using the same set of controls. We use a stacked design that combines all policy changes as instruments for a CAHOOTS response. Heteroskedasticity robust standard errors, clustered by call, are in parentheses. Significance stars indicate * p < 0.10, ** p < 0.05, *** p < 0.01

Table 4: Effects of CAHOOTS on call response and arrest using availability design

	Call responses			Call outcomes			
	CAHOOTS (1)	Police (2)	Any response (3)	Arrest (4)	Invol. hold (5)	Detox (6)	Warrant (7)
Panel A: Reduced form effect of high CAHOOTS capacity							
High CAHOOTS capacity (=1)	0.030***	-0.005**	0.016***	-0.001	0.000	0.000*	-0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)
Untreated mean	0.082	0.601	0.651	0.085	0.006	0.001	0.016
Panel B: IV estimates of effect of CAHOOTS response							
CAHOOTS		-0.173**	0.541***	-0.037	0.000	0.010*	-0.021
		(0.071)	(0.068)	(0.041)	(0.011)	(0.005)	(0.018)
Control complier mean		0.459	0.459	0.077	0.024	-0.004	0.022
First-stage F-stat		525.73	525.73	525.73	525.73	525.73	525.73
Hausman test							
χ^2 statistic		6.534	15.319	0.234	1.063	2.819	0.026
p-value		0.011	0.000	0.628	0.303	0.093	0.873
Equality with expansion effect (p)		0.042	0.000	0.006	0.000	0.000	0.162
Panel C: OLS estimates of effect of CAHOOTS response							
CAHOOTS		-0.239***	0.384***	-0.034***	0.012***	0.001***	-0.012***
		(0.004)	(0.002)	(0.002)	(0.001)	(0.000)	(0.001)
Untreated mean		0.620	0.620	0.087	0.005	0.001	0.017
Observations		238,333	238,333	238,333	238,333	238,333	238,333

Notes: This table presents estimates of the impact of a CAHOOTS response on call responses and call outcomes using our availability design. Panel A reports reduced form estimates of the effect of an indicator for having high CAHOOTS capacity when the focal call comes on these call outcomes. Panel B shows 2SLS estimates of the effect of a CAHOOTS response on call outcomes using the CAHOOTS capacity indicator as an instrument for having a CAHOOTS response. Panel C shows the association between a CAHOOTS response and these outcomes using the same set of controls. Heteroskedasticity robust standard errors are in parentheses. Significance stars indicate * p < 0.10, ** p < 0.05, *** p < 0.01

Table 5: Potential response types for a binary instrument that shifts CAHOOTS availability

		More availability			
		No response (n)	Only CAHOOTS (c)	Only police (p)	Both (b)
Less availability	No response (n)	π_{nn}	π_{nc}	-	-
	Only CAHOOTS (c)	-	π_{cc}	-	-
	Only police (p)	-	π_{pc}	π_{pp}	π_{pb}
	Both (b)	-	-	-	π_{bb}

Notes: This table presents the potential response types allowed under our choice model, where $\pi_{xy} \equiv Pr[D(0){=}d, D(1){=}d']$. The blue shading denotes the response types that are ruled out by Assumption 1 (monotonicity), and the orange shading denotes response types ruled out by Assumption 2 (no cross effects except substitution).

Table 6: Cross-design comparison of complier mean arrest rates

	Design	
	Availability	Diff-in-diff
	(1)	(2)
Panel A: Comparing $E[Y(b)]$ among $p{\to}b$ compliers		
$E[Y(b) \mid p{\to}b]$	0.138***	0.136***
	(0.032)	(0.015)
p-value of equality		0.948
J test of overidentification		
χ^2 statistic		0.595
p-value		0.964
Panel B: Comparing $E[Y(p)]$ among $p{\to}b$ and $p{\to}c$ compliers		
$\frac{s_{pb}}{s_{pc}+s_{pb}}E[Y(p) \mid p{\to}b] + \frac{s_{pc}}{s_{pc}+s_{pb}}E[Y(p) \mid p{\to}c]$	0.169***	0.315***
	(0.081)	(0.063)
p-value of equality		0.154
$s_{pb}/(s_{pc}+s_{pb})$	0.623	0.590
$s_{pc}/(s_{pc}+s_{pb})$	0.377	0.410

Notes: This table presents estimates of complier mean arrest rates in each of the research designs. Panel A shows an estimate of $E[Y(b)|p \to b]$ using the first moment in Equation 7. Panel B shows the weights and the estimated moment in the second equation in Equation 7. Significance stars indicate * p $<$ 0.10, ** p $<$ 0.05, *** p $<$ 0.01

Table 7: Complier group-specific treatment effects and outcome means for arrest

	(1)
Panel A: Complier means	
$E[Y(b) \mid p{\to}b]$	0.133***
	(0.013)
$E[Y(p) \mid p{\to}b]$	0.496***
	(0.163)
$E[Y(p) \mid p{\to}c]$	0.089
	(0.189)
Panel B: subLATEs	
Police $\to$ both	-0.363**
	(0.164)
Police $\to$ CAHOOTS	-0.0890
	(0.189)
Panel C: Goodness-of-fit tests	
J test of overidentification	
χ^2 statistic	1.032
p-value	0.994
J test of overidentification for $E[Y(p)]$ moment	
χ^2 statistic	0.438
p-value	0.993

Notes: This table reports estimates of the complier means and associated treatment effects. Our approach is to the complier means $E[Y(d) \mid g]$ that best fit the moments

$$\frac{E[Y1[D{=}b \mid Z{=}1]] - E[Y1[D{=}b \mid Z{=}0]]}{E[1[D{=}b \mid Z{=}1]] - E[1[D{=}b \mid Z{=}0]]} = E[Y(b) \mid p{\to}b]$$

$$\frac{E[Y1[D{=}p \mid Z{=}1]] - E[Y1[D{=}p \mid Z{=}0]]}{E[1[D{=}p \mid Z{=}1]] - E[1[D{=}p \mid Z{=}0]]} = \frac{s_{pb}}{s_{pb}+s_{pc}} E[Y(p) \mid p{\to}b] + \frac{s_{pb}}{s_{pb}+s_{pc}} E[Y(p) \mid p{\to}c].$$

Since there are five expansions, we have an overidentified set of equations with ten moments and three unknowns. We variance-weight the moments using a diagonal weighting matrix and solve via minimum distance. We also report J tests of overidentification for all moments together, as well as for the second moment (the J test for the $E[Y(b) \mid p{\to}b]$ moment can be found in Table 6). Significance stars indicate * p $<$ 0.10, ** p $<$ 0.05, *** p $<$ 0.01

Table 8: Effects of CAHOOTS on access to medical services

	Any Fire/ EMS response (1)	Suicidal person (2)	Overdose (3)	Cardiac arrest (4)	Transport to ER (5)	Dead on arrival (6)
Panel A: IV estimates of CAHOOTS response instrumented by CAHOOTS expansions						
CAHOOTS	0.26***	-0.037	0.031	0.044**	0.088**	-0.0029
	(0.061)	(0.029)	(0.022)	(0.022)	(0.037)	(0.006)
Control complier mean	-0.089	0.185	-0.012	-0.022	-0.049	0.005
	(.059)	(.026)	(.021)	(.022)	(.036)	(.006)
Observations	293,845	293,845	293,845	293,845	293,845	293,845
Panel B: OLS estimates with expansion design controls						
CAHOOTS	-0.01***	0.11***	0.0033***	0.0012	-0.013***	0.00047**
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Control mean	0.116	0.020	0.013	0.016	0.046	0.001
Observations	293,845	293,845	293,845	293,845	293,845	293,845
Panel C: IV estimates of CAHOOTS response instrumented by CAHOOTS availability						
CAHOOTS	-0.037	0.047**	-0.0053	0.000094	0.032	0.0013
	(0.046)	(0.022)	(0.017)	(0.015)	(0.025)	(0.003)
Control complier mean	0.128	0.057	0.030	-0.005	-0.007	-0.003
	(.043)	(.015)	(.016)	(.014)	(.024)	(.003)
Observations	238,333	238,333	238,333	238,333	238,333	238,333
Panel D: OLS estimates with availability design controls						
CAHOOTS	-0.0081***	0.12***	0.0039***	0.0011	-0.012***	0.00054***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Control mean	0.107	0.011	0.014	0.011	0.032	0.000
Observations	238,333	238,333	238,333	238,333	238,333	238,333

Notes: This table presents estimates of the impact of a CAHOOTS response on access to medical services. Panel A shows 2SLS estimates of the effect of a CAHOOTS response on call outcomes using indicators for each expansion as instruments. Panel B shows the association between a CAHOOTS response and these outcomes using the same stacked sample and set of controls. Panel C shows 2SLS estimates of the effect of a CAHOOTS response on call outcomes using the CAHOOTS capacity indicator as an instrument for having a CAHOOTS response. Panel D shows the association between a CAHOOTS response and these outcomes using the same sample and set of controls. Heteroskedasticity robust standard errors, clustered by call in Panels A and B, are in parentheses. Significance stars indicate * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

Adger, C., M. Ross, and C. Sloan (2025). The effect of field training officers on police use of force. Working paper.

Amaral, S., G. B. Dahl, V. Endl-Geyer, T. Hener, and H. Rainer (2023). Deterrence or backlash? arrests and the dynamics of domestic violence. Technical report, National Bureau of Economic Research.

Ang, D., P. Bencsik, J. M. Bruhn, and E. Derenoncourt (2024). Community engagement with law enforcemetn after high-profile acts of police violence. *NBER Working Paper*.

Anwar, S. and H. Fang (2006). An alternative test of racial prejudice in motor vehicle searches: Theory and evidence. *American Economic Review 96*(1), 127–151.

Ba, B., M. Baskar, T. Cheng, and R. Mariman (2024). Understanding demand for police alternatives. Working Paper.

Ba, B. A., P. J. Bayer, N. Rim, R. Rivera, and M. Sidibé (2021, September). Police officer assignment and neighborhood crime. Working Paper 29243, National Bureau of Economic Research.

Ba, B. A., D. Knox, J. Mummolo, and R. Rivera (2021). The role of officer race and gender in police-civilian interactions in chicago. *Science 371*(6530), 696–702.

Ba, B. A., N. Rim, and R. G. Rivera (2024). The relationship between officer misconduct and conviction-less arrests. Working paper, National Bureau of Economic Research.

Bazzi, S., R. A. Blair, C. Blattman, O. Dube, M. Gudgeon, and R. Peck (2022). The promise and pitfalls of conflict prediction: evidence from colombia and indonesia. *Review of Economics and Statistics 104*(4), 764–779.

Blais, E., M. Landry, N. Elazhary, S. Carrier, and A.-M. Savard (2022). Assessing the capability of a co-responding police-mental health program to connect emotionally disturbed people with community resources and decrease police use-of-force. *Journal of Experimental Criminology 18*, 41–65.

Blanes i Vidal, J. and T. Kirchmaier (2018). The effect of police response time on crime clearance rates. *The Review of Economic Studies 85*(2), 855–891.

Blattman, C., D. P. Green, D. Ortega, and S. Tobón (2021). Place-based interventions at scale: The direct and spillover effects of policing and city services on crime. *Journal of the European Economic Association 19*, 2022–2051.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, rdae007.

Chalfin, A. and F. M. Gonçalves (2023). Professional motivations in the public sector: Evidence from police officers. Technical report, National Bureau of Economic Research. Available at https://www.nber.org/papers/w31985.

Chalfin, A., B. Hansen, E. K. Weisburst, and M. C. Williams Jr (2022). Police force size and civilian race. *American Economic Review: Insights 4*(2), 139–158.

Chalfin, A. and J. McCrary (2018). Are us cities underpoliced? theory and evidence. *Review of Economics and Statistics 100*(1), 167–186.

Cohen, G. and M. Bagwell (2023). The state of mental health training in basic police curricula: A national level examination. *Journal of Public Affairs Education 29*(3), 327–349.

Cohen, M. A. and A. R. Piquero (2009). New evidence on the monetary value of saving a high risk youth. *Journal of Quantitative Criminology 25*, 25–49.

Dee, T. S. and J. Pyne (2022). A community response approach to mental health and substance abuse crises reduced crime. *Science Advances 8*(23), 1–9.

Di Tella, R. and E. Schargrodsky (2004). Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. *American Economic Review 94*(1), 115–133.

Draca, M., S. Machin, and R. Witt (2011). Panic on the streets of london: Police, crime, and the july 2005 terror attacks. *American Economic Review 101*(5), 2157–81.

Eugene Register Guard (2024). Cahoots and hoots workers seek fair contract with white bird clinic. Available at https://www.registerguard.com/story/news/local/2024/04/06/cahoots-and-hoots-workers-seek-fair-contract-with-white-bird-clinic-in-eugene-garners 73225506007/.

Evans, W. N. and E. G. Owens (2007). Cops and crime. *Journal of Public Economics 91*(1-2), 181–201.

Feigenberg, B. and C. Miller (2022). Would eliminating racial disparities in motor vehicle searches have efficiency costs? *The Quarterly Journal of Economics 137*(1), 49–113.

Feigenberg, B. and C. Miller (2023). Class disparities and discrimination in traffic stops and searches. Technical report, Working paper.

Finkelstein, A. and N. Hendren (2020). Welfare analysis meets causal inference. *Journal of Economic Perspectives 34*(4), 146–167.

Gagnon, L. R., L. Reetz, and S. Radcliffe (2022). Mental health holds and civil commitment. Available at https://www.droregon.org/disability-rights-oregon-resources/mental-health-holds-and-civil-commitment.

Goncalves, F. and S. Mello (2021). A few bad apples? racial bias in policing. *American Economic Review 111*(5), 1406–1441.

Heckman, J. J. and R. Pinto (2018). Unordered monotonicity. *Econometrica 86*(1), 1–35.

Helfgott, J. B., M. J. Hickman, and A. P. Labossiere (2016). A descriptive evaluation of the seattle police department's crisis response team officer/mental health professional partnership pilot program. *International Journal of Law and Psychiatry 44*, 109–122.

Heslin, M., L. Callaghan, M. Packwood, V. Badu, and S. Byford (2016). Decision analytic model exploring the cost and cost-offset implications of street triage. *BMJ Open 6*.

Hoekstra, M. and C. Sloan (2022). Does race matter for police use of force? evidence from 911 calls. *American Economic Review 112*, 827–860.

IACP (2018). Responding to persons experiencing a mental health crisis. Available at https://www.theiacp.org/sites/default/files/2018-08/MentalIllnessBinder2018.pdf.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

Irwin, A. and B. Pearl (2020). The community responder model: How cities can send the right responder to every 911 call. Available at https://www.americanprogress.org/article/community-responder-model/.

Jenkins, O., S. Dye, F. Obeng-Asare, N. Nguyen, and N. Wright (2017). Police liaison and section 136: comparison of two different approaches. *BJPsych Bull 41*, 76–82.

Keown, P., J. French, G. Gibson, E. Newton, S. Cull, P. Brown, J. Parry, D. Lyons, and I. McKinnon (2016). Too much detention? street triage and detentions under section 136 mental health act in the north-east of england: a descriptive study of the effects of a street triage intervention. *BMJ Open 6*.

Klick, J. and A. Tabarrok (2005). Using terror alert levels to estimate the effect of police on crime. *The Journal of Law and Economics 48*(1), 267–279.

Kline, P. and C. R. Walters (2016). Evaluating public programs with close substitutes: The case of head start. *The Quarterly Journal of Economics 131*(4), 1795–1848.

Krider, A., R. Huerter, K. Gaherty, and A. Moore (2020). Responding to individuals in behavioral health crisis via co-responder models: The roles of cities, counties, law enforcement, and providers. Available at https://www.prainc.com/wp-content/uploads/2020/03/RespondingtoBHCrisisviaCRModels.pdf.

Lum, C. (2021). Perspectives on policing: Cynthia lum. *Annual Review of Criminology 4*(1), 19–25.

Marcus, N. and V. Stergiopoulos (2022). Re-examining mental health crisis intervention: A rapid review comparing outcomes across police, co-responder and non-police models. *Health & social care in the community 30*(5), 1665–1679.

Mello, S. (2019). More cops, less crime. *Journal of Public Economics 172*, 174–200.

Mello, S. (2024). Fines and financial wellbeing. *Review of Economic Studies*, rdae111.

Mikdash, M. and R. Zaiour (2024). The impact of police shootings on gun violence and civilian cooperation. *Journal of Public Economics 237*, 105189.

Miller, A. R. and C. Segal (2019). Do female officers improve law enforcement quality? effects on crime reporting and domestic violence. *The Review of Economic Studies 86*(5), 2220–2247.

Morabito, M. S., J. Savage, L. Sneider, and K. Wallace (2018). Police response to people with mental illnesses in a major us city: The boston experience with the co-responder model. *Victims & Offenders 13*(8), 1093–1105.

Moreno-Medina, J., A. Ouss, P. Bayer, and B. A. Ba (2023). Officer-involved: The media language of police killings. Technical report, Accepted *Quarterly Journal of Economics*. Available at https://static1.squarespace.com/static/5ec5aa4736bba52179f2ff85/t/65dcf3227d4dbf14c32d1f3a/1708978979966/police_language.pdf.

Owens, E. (2019). An economic approach to de-policing. *Criminology & Public Policy 18*(1), 77–80.

Premkumar, D. (2019, sep). Public scrutiny, police behavior, and crime consequences: Evidence from high-profile police killings.

Rivera, R. (2025a). Are 'bad' cops better police? the trade-off between officer aggression and public safety. Working paper.

Rivera, R. (2025b, April). Do peers matter in the police academy? *American Economic Journal: Applied Economics 17*(2), 127–164.

Shapiro, G. K., A. Cusi, M. Kirst, P. O'Campo, A. Nakhost, and V. Stergiopoulos (2015). Co-responding police-mental health programs: A review. *Administration and policy in mental health and mental health services research 42*, 606–620.

Walker, S., E. Mackay, P. Barnett, L. S. Rains, M. Leverton, C. Dalton-Locke, K. Trevillion, B. Lloyd-Evans, and S. Johnson (2019). Clinical and social factors associated with increased risk for involuntary psychiatric hospitalisation: a systematic review, meta-analysis, and narrative synthesis. *The Lancet Psychiatry 6*(12), 1039–1053.

Weisburst, E. K. (2019). Safety in police numbers: Evidence of police effectiveness from federal cops grant applications. *American Law and Economics Review 21*, 81–109.

White Bird Clinic (2020). *CAHOOTS Media Guide*. Accessed April 1, 2025.

A1 Econometric Appendix

Given the four treatments, there are eight moments that reflect the effect of the instrument on the outcome multiplied by an indicator for receiving or not receiving a treatment d, rescaled by the relevant first stage. These moments are as follows:

$$\begin{bmatrix} \frac{E[Y1[D=b|Z=1]]-E[Y1[D=b|Z=0]]}{E[1[D=b|Z=1]]-E[1[D=b|Z=0]]} \\ \frac{E[Y1[D\neq b|Z=1]]-E[Y1[D\neq b|Z=0]]}{E[1[D\neq b|Z=1]]-E[1[D\neq b|Z=0]]} \\ \frac{E[Y1[D=c|Z=1]]-E[Y1[D=c|Z=0]]}{E[1[D=c|Z=1]]-E[1[D=c|Z=0]]} \\ \frac{E[Y1[D\neq c|Z=1]]-E[Y1[D\neq c|Z=0]]}{E[1[D\neq c|Z=1]]-E[1[D\neq c|Z=0]]} \\ \frac{E[Y1[D=p|Z=1]]-E[Y1[D=p|Z=0]]}{E[1[D=p|Z=1]]-E[1[D=p|Z=0]]} \\ \frac{E[Y1[D\neq p|Z=1]]-E[Y1[D\neq p|Z=0]]}{E[1[D\neq p|Z=1]]-E[1[D\neq p|Z=0]]} \\ \frac{E[Y1[D=n|Z=1]]-E[Y1[D=n|Z=0]]}{E[1[D=n|Z=1]]-E[1[D=n|Z=0]]} \\ \frac{E[Y1[D\neq n|Z=1]]-E[Y1[D\neq n|Z=0]]}{E[1[D\neq n|Z=1]]-E[1[D\neq n|Z=0]]} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & \frac{-\pi_{pc}}{\pi_{pb}} & \frac{\pi_{pc}}{\pi_{pb}} & \frac{-\pi_{nc}}{\pi_{pb}} & \frac{\pi_{nc}}{\pi_{pb}} \\ 0 & 0 & \frac{\pi_{pc}}{\pi_{pc}+\pi_{nc}} & 0 & \frac{\pi_{nc}}{\pi_{pc}+\pi_{nc}} & 0 \\ \frac{-\pi_{pb}}{\pi_{pc}+\pi_{nc}} & \frac{\pi_{pb}}{\pi_{pc}+\pi_{nc}} & 0 & \frac{\pi_{pc}}{\pi_{pc}+\pi_{nc}} & 0 & \frac{\pi_{nc}}{\pi_{pc}+\pi_{nc}} \\ 0 & \frac{\pi_{pb}}{\pi_{pb}+\pi_{pc}} & 0 & \frac{\pi_{pc}}{\pi_{pb}+\pi_{pc}} & 0 & 0 \\ \frac{\pi_{pb}}{\pi_{pb}+\pi_{pc}} & 0 & \frac{\pi_{pc}}{\pi_{pb}+\pi_{pc}} & 0 & \frac{\pi_{nc}}{\pi_{pb}+\pi_{pc}} & \frac{-\pi_{nc}}{\pi_{pb}+\pi_{pc}} \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{\pi_{pb}}{\pi_{nc}} & \frac{-\pi_{pb}}{\pi_{nc}} & \frac{\pi_{pc}}{\pi_{nc}} & \frac{-\pi_{pc}}{\pi_{nc}} & \frac{\pi_{nc}}{\pi_{nc}} & 0 \end{bmatrix} \begin{bmatrix} E[Y(b) \mid p\to b] \\ E[Y(p) \mid p\to b] \\ E[Y(c) \mid p\to c] \\ E[Y(p) \mid p\to c] \\ E[Y(c) \mid n\to c] \\ E[Y(n) \mid n\to c] \end{bmatrix}$$

Note that:

$$\frac{E[Y1[D=d] \cdot 1[Z=z]]|Z=1] - E[Y1[D=d] \cdot 1[Z=z]|Z=0]}{E[1[D=d]|Z=1] - E[1[D=d]|Z=0]} = E[Y|D=d, Z=z]$$

and therefore the moments we express above in the form of Wald estimators can be mapped to $E[Y|D=d, Z=z]$.

Given that in our setting, $E[Y(c) \mid p\to c] = E[Y(c) \mid n\to c] = E[Y(n) \mid n\to c] = 0$, this expression can be simplified to

$$\begin{bmatrix} \frac{E[Y1[D=b|Z=1]]-E[Y1[D=b|Z=0]]}{E[1[D=b|Z=1]]-E[1[D=b|Z=0]]} \\ \frac{E[Y1[D\neq b|Z=1]]-E[Y1[D\neq b|Z=0]]}{E[1[D\neq b|Z=1]]-E[1[D\neq b|Z=0]]} \\ \frac{E[Y1[D=c|Z=1]]-E[Y1[D=c|Z=0]]}{E[1[D=c|Z=1]]-E[1[D=c|Z=0]]} \\ \frac{E[Y1[D\neq c|Z=1]]-E[Y1[D\neq c|Z=0]]}{E[1[D\neq c|Z=1]]-E[1[D\neq c|Z=0]]} \\ \frac{E[Y1[D=p|Z=1]]-E[Y1[D=p|Z=0]]}{E[1[D=p|Z=1]]-E[1[D=p|Z=0]]} \\ \frac{E[Y1[D\neq p|Z=1]]-E[Y1[D\neq p|Z=0]]}{E[1[D\neq p|Z=1]]-E[1[D\neq p|Z=0]]} \\ \frac{E[Y1[D=n|Z=1]]-E[Y1[D=n|Z=0]]}{E[1[D=n|Z=1]]-E[1[D=n|Z=0]]} \\ \frac{E[Y1[D\neq n|Z=1]]-E[Y1[D\neq n|Z=0]]}{E[1[D\neq n|Z=1]]-E[1[D\neq n|Z=0]]} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{\pi_{pc}}{\pi_{pb}} \\ 0 & 0 & 0 \\ \frac{-\pi_{pb}}{\pi_{pc}+\pi_{nc}} & \frac{\pi_{pb}}{\pi_{pc}+\pi_{nc}} & \frac{\pi_{pc}}{\pi_{pc}+\pi_{nc}} \\ 0 & \frac{\pi_{pb}}{\pi_{pb}+\pi_{pc}} & \frac{\pi_{pc}}{\pi_{pb}+\pi_{pc}} \\ \frac{\pi_{pb}}{\pi_{pb}+\pi_{pc}} & 0 & 0 \\ 0 & 0 & 0 \\ \frac{\pi_{pb}}{\pi_{nc}} & \frac{-\pi_{pb}}{\pi_{nc}} & \frac{-\pi_{pc}}{\pi_{nc}} \end{bmatrix} \begin{bmatrix} E[Y(b) \mid p\to b] \\ E[Y(p) \mid p\to b] \\ E[Y(p) \mid p\to c] \end{bmatrix}$$

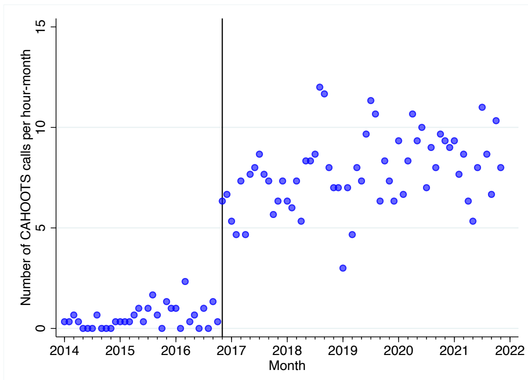
Further simplifications are possible. The third and seventh line can be removed because the weight is zero on each potential outcome, the second and fifth line are collinear because the arrest outcome is necessarily zero whenever $D \in \{n, c\}$, the fourth line is a linear combination of the first and fifth line, the sixth line is a rescaled version of the first line, and the eight line is colinear with the fourth. This leaves only:

$$\begin{bmatrix} \frac{E[Y1[D=b|Z=1]]-E[Y1[D=b|Z=0]]}{E[1[D=b|Z=1]]-E[1[D=b|Z=0]]} \\ \frac{E[Y1[D=p|Z=1]]-E[Y1[D=p|Z=0]]}{E[1[D=p|Z=1]]-E[1[D=p|Z=0]]} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\pi_{pb}}{\pi_{pb}+\pi_{pc}} & \frac{\pi_{pc}}{\pi_{pb}+\pi_{pc}} \end{bmatrix} \begin{bmatrix} E[Y(b) \mid p\to b] \\ E[Y(p) \mid p\to b] \\ E[Y(p) \mid p\to c] \end{bmatrix} \tag{A1}$$

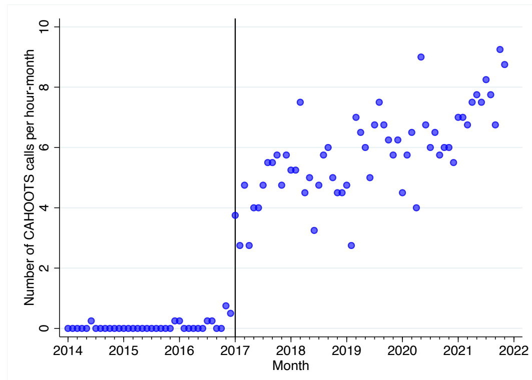
A2 Appendix Figures

Figure A1: CAHOOTS expansions in Eugene and Springfield, 911 calls only

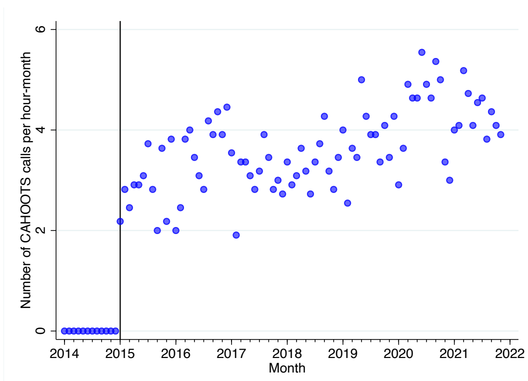
(a) Eugene 7-10 a.m.



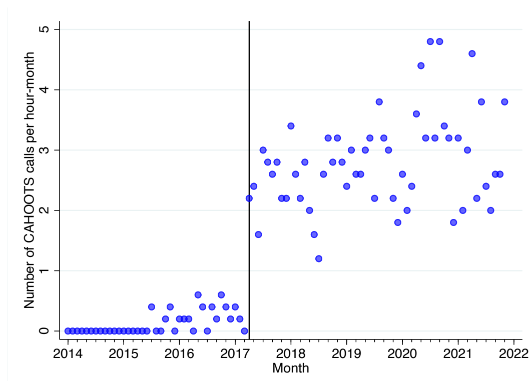
(b) Eugene 3-7 a.m.



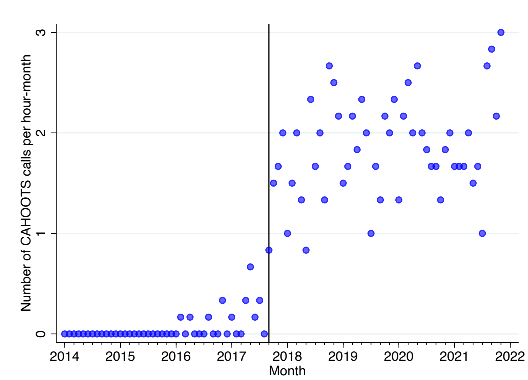
(c) Springfield 12 p.m. to 11 p.m.



(d) Springfield 9-11 a.m. and 11 p.m.-2 a.m.



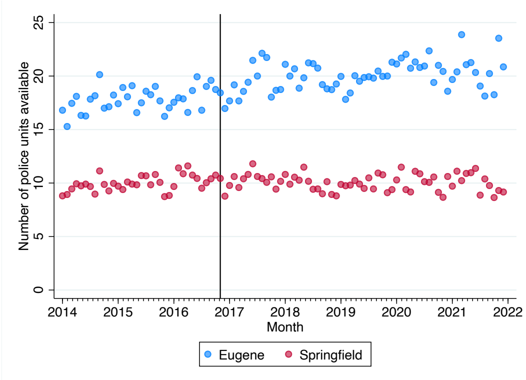
(e) Springfield 3-10 a.m.



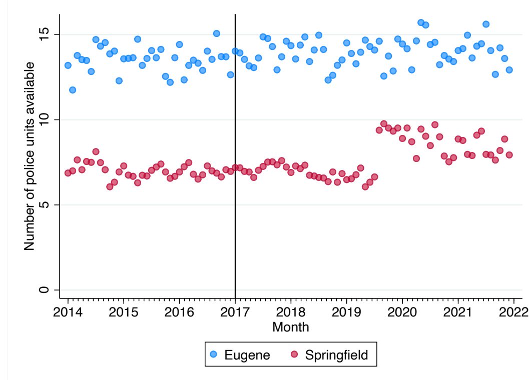
Notes: This figure shows the changes in the number of calls to which CAHOOTS is dispatched as services are expanded (Eugene) or introduced (Springfield). Panels A and B report expansions in Eugene and Panels C, D, and E expansions in Springfield. Each point in the figures reports the number of 911 calls to which CAHOOTS was dispatched, normalized to be measured in terms of calls per hour-month.

Figure A2: Police units available during design hours

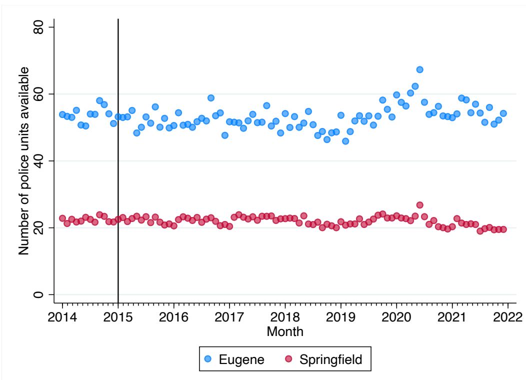
(a) 7-10 a.m.



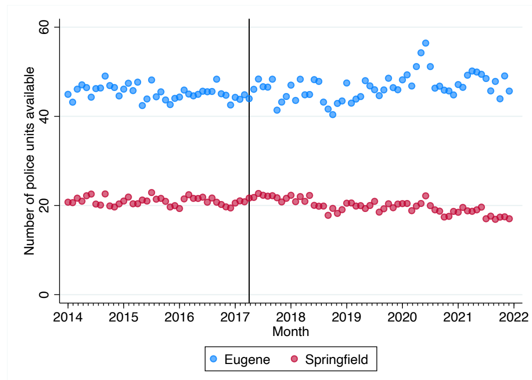
(b) 3-7 a.m.



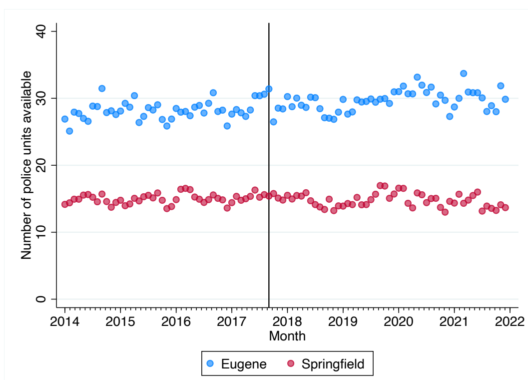
(c) 12 p.m. to 11 p.m.



(d) 9-11 a.m. and 11 p.m.-2 a.m.

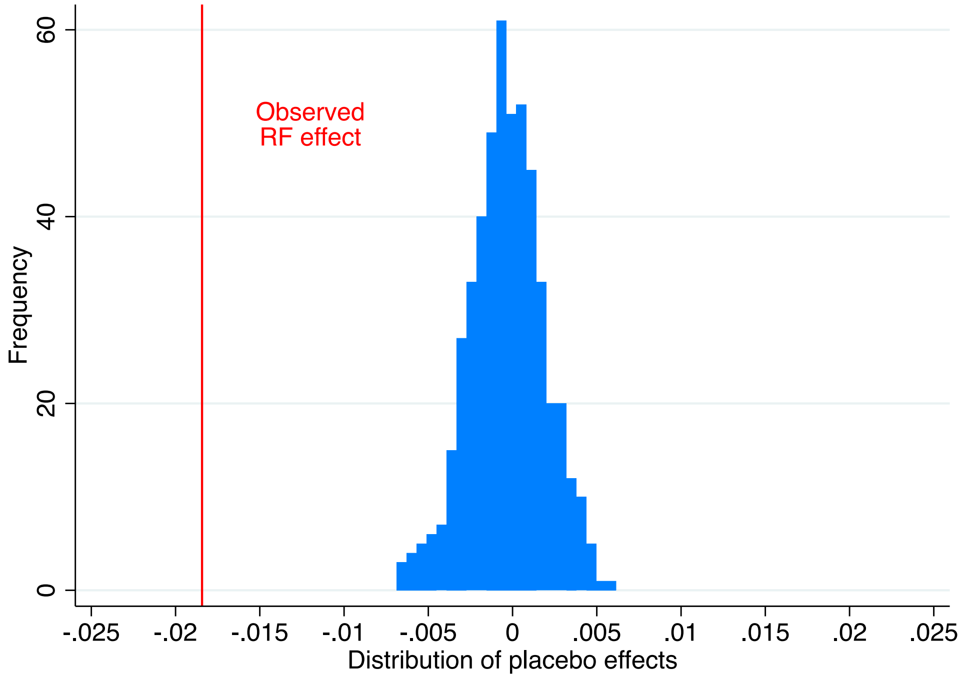


(e) 3-10 a.m.



Notes: This figure shows the average number of police units available during design hours in each month and city.

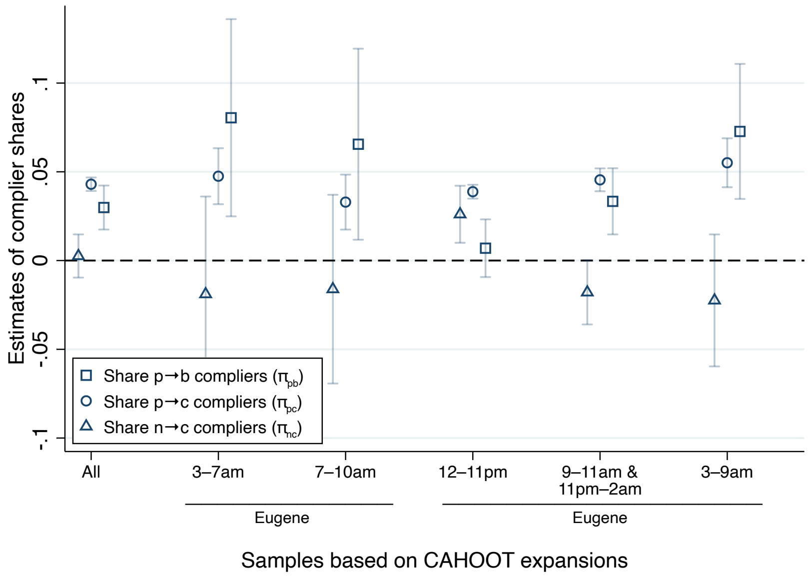
Figure A3: Distribution of placebo expansion times



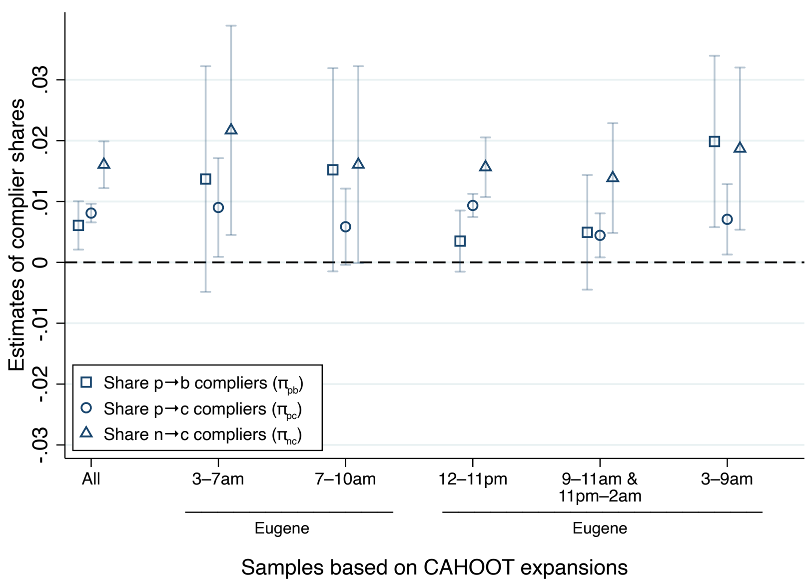
Notes: This figure displays the distribution of estimated effects for 500 randomly chosen event times. These effects are estimated using the same specification as in Equation 3, focusing on the reduced-form relationship. The placebo event time effects are small, centered around zero, and do not overlap with the estimated effect from the observed data.

Figure A4: Tests of model assumptions

(a) Expansions design



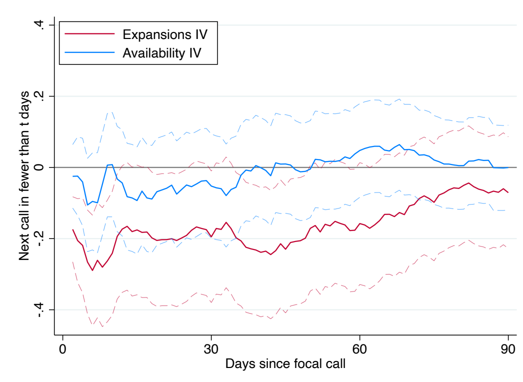
(b) Availability design



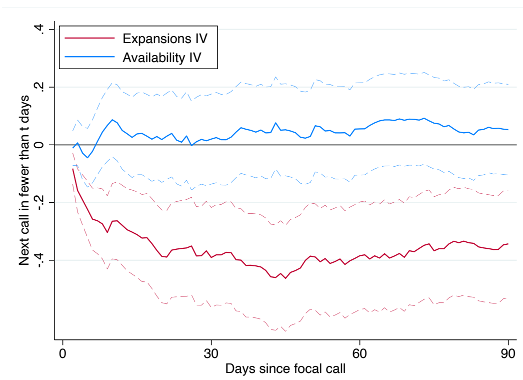
Notes: The figure reports estimates of the share of the three complier types: π_{pb}, π_{pc}, and π_{nc} across different sub-samples for each design. The x-axis shows the sub-sample, and the y-axis shows the estimated share of compliers. The sub-samples are defined based on the data that is used to estimate each specific expansion. Negative values of any of the estimated shares will indicate a potential violation of some of the model assumptions. The estimation is done by estimating Equation 5 in each sub-sample separately.

Figure A5: Robustness for analysis of effect of CAHOOTS response on future calls from area in fewer than t days

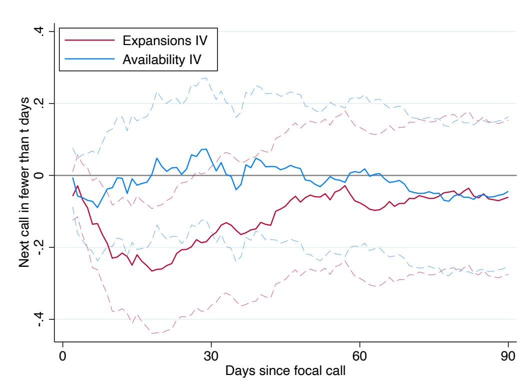
(a) Within 50 meters of focal call



(b) Within 50 meters of focal call, excluding exact location



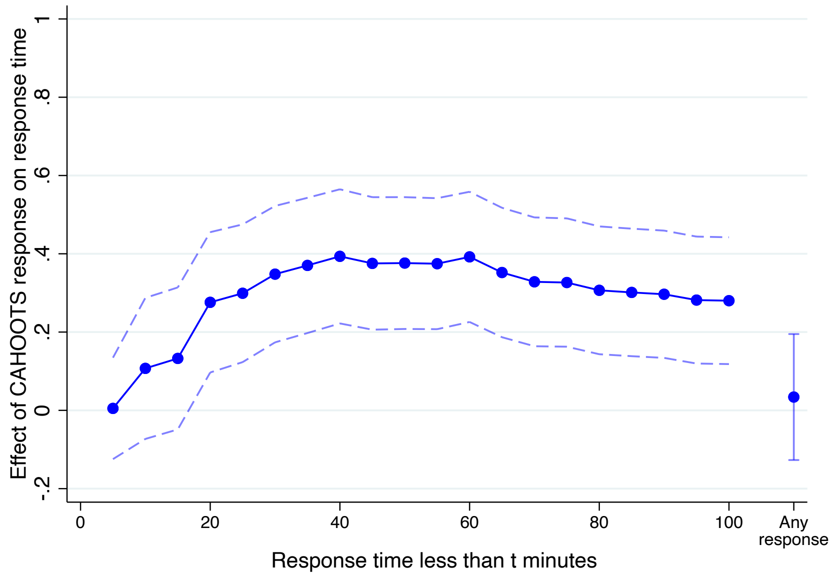
(c) Excluding high-traffic and public places



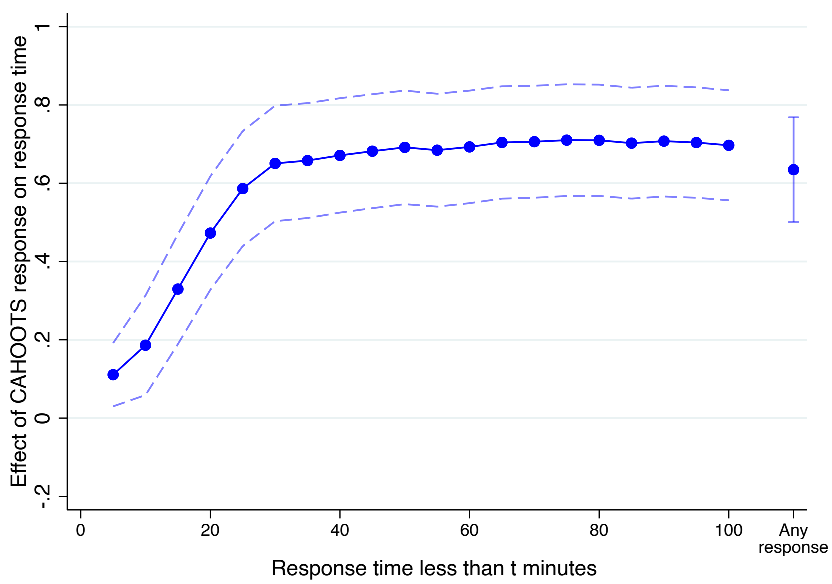
Notes: Panel A shows the effect of a CAHOOTS response on future 911 calls from within 50 meters of the focal call's location within a period of time up to 90 days. Panel B shows analogous results, excluding the exact location of the focal call. Panel C excludes locations with more than 300 calls or with a non-numbered address such as an intersection of a block of a street. Each coefficient comes from a 2SLS regression of an indicator for a future 911 call within t days on a CAHOOTS response, where CAHOOTS response is instrumented using either indicators for each CAHOOTS expansion or the availability instrument indicator. 95% confidence intervals clustered at the individual level are shown in dotted lines.

Figure A6: Effect of CAHOOTS response on response in less than t minutes

(a) Expansions design



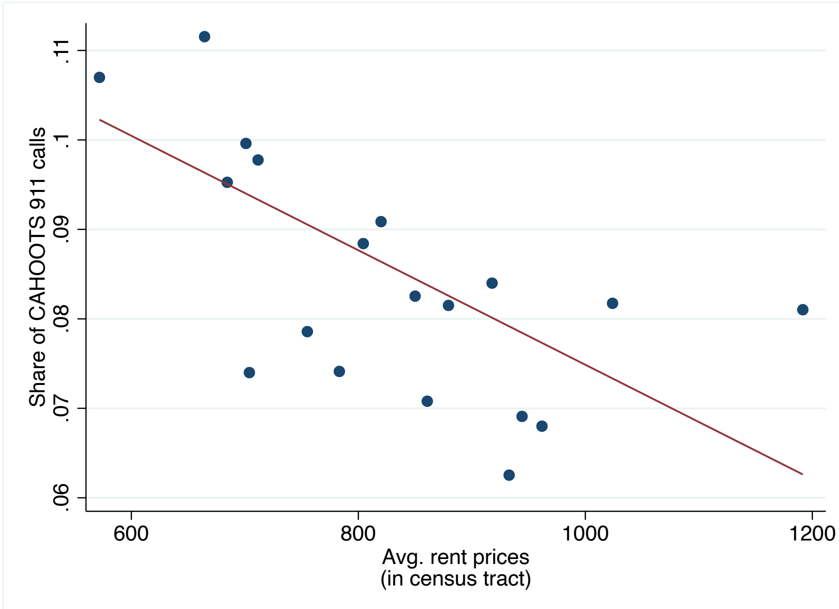
(b) Busyness design



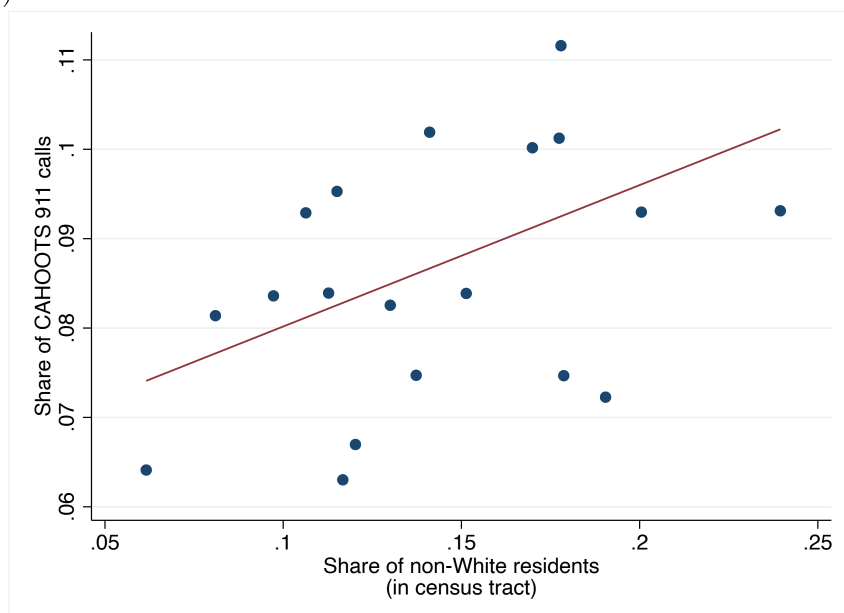
Notes: This figure shows the effect of a CAHOOTS response on time to response. Panel A shows the impact of CAHOOT responses induced by the expansions design and Panel B the affects of responses induced by the availability design. Each dot represents the coefficient from single 2SLS regression where the outcome is a response occurring in less than t minutes.

Figure A7: The correlation between CAHOOTS services and neighborhood characteristics

(a) CAHOOTS share of 911 calls and share of non-White residents



(b) CAHOOTS share of 911 calls and share of non-White residents



Notes: The neighborhood characteristics are measured using the American Community Survey at the tract level and are merged to the 911 call data based on the geographic coordinates of the 911 call.

A3 Appendix Tables

Table A1: CAHOOTS expansions in Eugene and Springfield since 2014

	(1) Date	(2) Coverage	(3) Change
		Eugene	
Study baseline	As of 1/1/2014	11-3 a.m.	16
Expansion 1	11/1/2016	+ 7-10 a.m.	+4 (20 total)
Expansion 2	1/1/2017	+ 3-7 a.m	+4 (24 total)
		Springfield	
Study baseline	As of 1/1/2014	-	0
Expansion 3	1/14/2015	+ noon-11 p.m.	+11 (11 total)
Expansion 4	4/1/2017	+ 9-11 a.m. & 11 p.m.-2 a.m.	+7 (18 total)
Expansion 5	9/1/2017	+ 3-10 a.m.	+6 (24 total)

Notes: This table reports the start dates and resulting hour-of-day coverage of the five CAHOOTS expansions in Eugene and Springfield.

Table A2: Robustness analysis of effects of CAHOOTS using expansion design

	Call responses		Call outcomes			
	Police	Any response	Arrest	Invol. hold	Detox	Warrant
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: IV estimates of effect of CAHOOTS response						
CAHOOTS	-0.396***	0.034	-0.240***	-0.151***	-0.051***	0.023
	(0.084)	(0.082)	(0.061)	(0.023)	(0.013)	(0.026)
Control complier mean	0.966	0.966	0.317	0.172	0.054	-0.012
First-stage F-stat	177.76	177.76	177.76	177.76	177.76	177.76
Panel B: adding geographic tract controls						
CAHOOTS	-0.378***	0.055	-0.232***	-0.151***	-0.050***	0.025
	(0.083)	(0.081)	(0.061)	(0.023)	(0.013)	(0.025)
Control complier mean	0.945	0.945	0.309	0.171	0.054	-0.014
First-stage F-stat	180.48	180.48	180.48	180.48	180.48	180.48
Panel C: adding geographic tract and call nature controls						
CAHOOTS	-0.332***	0.105	-0.241***	-0.143***	-0.052***	-0.003
	(0.082)	(0.080)	(0.057)	(0.022)	(0.013)	(0.017)
Control complier mean	0.895	0.895	0.317	0.164	0.055	0.013
First-stage F-stat	170.35	170.35	170.35	170.35	170.35	170.35
Panel D: adding tract FEs						
CAHOOTS	-0.361***	0.071	-0.226***	-0.151***	-0.050***	0.027
	(0.083)	(0.081)	(0.061)	(0.023)	(0.013)	(0.026)
Control complier mean	0.929	0.929	0.303	0.172	0.053	-0.016
First-stage F-stat	178.07	178.07	178.07	178.07	178.07	178.07
Observations	293,845	293,845	293,845	293,845	293,845	293,845

Notes: This table explores the sensitivity of the estimates in Table 3 to the inclusion of additional controls. Each panel adds successively more controls. Robust standard errors, clustered by call, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Reduced form effects of each CAHOOTS expansion

	Call responses			Call outcomes			
	(1) CAHOOTS	(2) Police	(3) Any response	(4) Arrest	(5) Invol. hold	(6) Detox	(7) Warrant
Eugene 3-7 a.m., after Jan. 1 2017	0.11***	-0.080***	-0.019	-0.033	0.00034	-0.0024	-0.00083
	(0.011)	(0.028)	(0.028)	(0.023)	(0.0079)	(0.0047)	(0.011)
Eugene 7-10 a.m., after Nov. 1 2016	0.082***	-0.066**	-0.016	-0.0094	0.0077	-0.0046	-0.0064
	(0.011)	(0.027)	(0.027)	(0.019)	(0.0069)	(0.0038)	(0.0082)
Springfield noon-11 p.m., after Jan. 14 2015	0.072***	-0.0070	0.026***	-0.014**	-0.015***	-0.0041***	0.0057***
	(0.0031)	(0.0083)	(0.0082)	(0.0059)	(0.0024)	(0.0012)	(0.0022)
Springfield 9-11 a.m. and 11 p.m.-2 a.m., after Apr. 1 2017	0.061***	-0.033***	-0.018*	-0.021***	-0.014***	-0.0038***	0.00038
	(0.0047)	(0.0095)	(0.0092)	(0.0072)	(0.0026)	(0.0015)	(0.0031)
Springfield 3-9 a.m., after Sept. 1 2017	0.11***	-0.073***	-0.022	-0.025*	-0.0056	-0.0035	-0.0058
	(0.010)	(0.019)	(0.019)	(0.013)	(0.0040)	(0.0026)	(0.0061)
Variance-weighted avg. effect	0.092	-0.063	-0.015	-0.022	-0.002	-0.004	-0.003
	(0.006)	(0.015)	(0.015)	(0.011)	(0.004)	(0.002)	(0.005)
Observations	293,845	293,845	293,845	293,845	293,845	293,845	293,845

Notes: This table reports the effect of CAHOOTS expansions on types of call responses and call outcomes. Each column reports the coefficients from a regression of the outcome variable on indicators for each CAHOOTS expansion. City-design and month-design fixed effects and controls for other policy changes are included. Standard errors clustered at the call level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Traditional diff-in-diff and Borusyak, Javier, and Spiess (2024)

	Call responses			Call outcomes			
	CAHOOTS (1)	Police (2)	Any response (3)	Arrest (4)	Invol. hold (5)	Detox (6)	Warrant (7)
Panel A: Traditional diff-in-diff							
Post expansion	0.075***	-0.030***	0.003	-0.018***	-0.011***	-0.004***	0.002
	(0.003)	(0.006)	(0.006)	(0.005)	(0.002)	(0.001)	(0.002)
Dependent mean	0.081	0.599	0.651	0.085	0.007	0.002	0.016
Observations	293,845	293,845	293,845	293,845	293,845	293,845	293,845
Panel B: BJS imputation estimator							
Post expansion	0.076***	-0.039***	-0.002	-0.023***	-0.008**	-0.003*	0.003
	(0.003)	(0.011)	(0.011)	(0.008)	(0.003)	(0.002)	(0.003)
Dependent mean	0.005	0.721	0.724	0.120	0.019	0.005	0.017
Observations	38,849	38,849	38,849	38,849	38,849	38,849	38,849

Notes: This table compares the traditional difference in differences to the imputation approach of Borusyak, Javier, and Spiess (2024). The traditional diff-in-diff differs only from our main reduced form approach by using an indicator for any expansion rather than separate indicators for each expansion; it therefore includes call hour-city-design and call month-design FEs. The BJS specification uses only the ever-treated observations, and includes call hour and call month FEs. To match the implicit weights in our main 2SLS specification, all reduced-form regressions are variance reweighted at the design level by the ratio of the variance of the residualized first stages for CAHOOTS using each expansion and an indicator for any expansion, respectively. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: 2SLS estimates of CAHOOTS effects using availability design

	Call responses		Call outcomes			
	Police	Any response	Arrest	Invol. hold	Detox	Warrant
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: IV estimates of effect of CAHOOTS response						
CAHOOTS	-0.175**	0.539***	-0.039	0.001	0.009*	-0.022
	(0.071)	(0.068)	(0.041)	(0.011)	(0.005)	(0.018)
Control complier mean	0.366	0.366	0.057	0.023	-0.004	0.016
First-stage F-stat	526.00	526.00	526.00	526.00	526.00	526.00
Panel B: adding geographic tract controls						
CAHOOTS	-0.182**	0.532***	-0.042	0.000	0.009*	-0.022
	(0.071)	(0.068)	(0.041)	(0.011)	(0.005)	(0.018)
Control complier mean	0.373	0.373	0.060	0.024	-0.004	0.017
First-stage F-stat	522.16	522.16	522.16	522.16	522.16	522.16
Panel C: adding geographic tract and call nature controls						
CAHOOTS	-0.201***	0.514***	-0.049	-0.004	0.010*	-0.021
	(0.072)	(0.069)	(0.043)	(0.012)	(0.006)	(0.018)
Control complier mean	0.391	0.391	0.070	0.029	-0.004	0.017
First-stage F-stat	545.34	545.34	545.34	545.34	545.34	545.34
Panel D: adding tract FEs						
CAHOOTS	-0.185***	0.529***	-0.044	0.000	0.009*	-0.023
	(0.071)	(0.068)	(0.041)	(0.011)	(0.005)	(0.018)
Control complier mean	0.376	0.376	0.061	0.024	-0.003	0.018
First-stage F-stat	519.22	519.22	519.22	519.22	519.22	519.22
Observations	238,330	238,330	238,330	238,330	238,330	238,330

Notes: This table explores the sensitivity of the estimates in Table 4 to the inclusion of additional controls. Each panel successively adds more controls. Robust standard errors in parentheses. Significance stars indicate * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Different subLATE weights do not explain cross-design differences in treatment effects

	Design		
	Availability	Diff-in-diff	
	2SLS	2SLS	Min. dist.
	(1)	(2)	(3)
Panel A: Treatment effect estimates			
Rescaled CAHOOTS effect ($\frac{s_{pb}}{s_{pc}+s_{pb}}\Delta_{pb}+\frac{s_{pc}}{s_{pc}+s_{pb}}\Delta_{pc}$)	-0.0857	-0.241	-0.241
	(0.0861)	(0.0616)	(0.0616)
Panel B: Weights			
$s_{pb}/(s_{pc}+s_{pb})$	0.623	0.590	0.590
$s_{pc}/(s_{pc}+s_{pb})$	0.377	0.410	0.410
Panel C: Predicted availability design LATEs using subLATEs			
Rescaled effect ($\frac{s_{pb}^{busy}}{s_{pc}^{busy}+s_{pb}^{busy}}\Delta_{pb}+\frac{s_{pc}^{busy}}{s_{pc}^{busy}+s_{pb}^{busy}}\Delta_{pc}$)		-0.263	-0.260
		(0.065)	(0.062)

Notes: This table reports estimates of the rescaled effect from Equation 6. Panel A shows the rescaled effects using the two research designs. For the difference-in-differences design, we also show a version estimating the subLATEs using minimum distance instead of IV. Panel B shows the weights on the subLATES. Panel C applies the availability design weights to the subLATE estimates from the expansion design to predict the availability design effect. Heteroskedasticity robust standard errors, clustered by call in columns (2) and (3), are in parentheses. Significance stars indicate * p < 0.10, ** p < 0.05, *** p < 0.01

Table A7: Sensitivity of estimated CAHOOTS effect on arrests to police busyness

	Diff-in-diff		Availability	
	(1)	(2)	(3)	(4)
CAHOOTS	-0.240***	-0.279***	-0.0167	-0.0381
	(0.0614)	(0.0636)	(0.0405)	(0.0406)
Control complier mean	0.317	0.317	0.078	0.078
First-stage F-stat	177.76	164.78	528.08	525.37
Observations	293,844	293,844	237,347	237,347
Control for police busyness	No	Yes	No	Yes

Notes: This table shows the sensitivity of the estimated impacts of dispatching CAHOOTS on arrests to controlling for police busyness using the differences-in-differences and availability designs. Heteroskedasticity robust standard errors, clustered by call in columns (1) and (2), are in parentheses. Significance stars indicate * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$