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

March 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 evaluate 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 traditional 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 a marginal response. We find that marginal additions to the program's capacity 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, shemtov@econ.ucla.edu, and cstrickland@tamu.edu, respectively. We thank Bocar Ba, Felipe Goncalves, Jens Ludwig, Grant McDermott, Marit Rehavi, Evan K. Rose, and seminar participants at multiple universities for their helpful comments. Diti Jain and Everett Stamm provided outstanding research assistance.

Police officers respond to a broad range of 911 calls, spanning all the way from violent crimes in progress to non-criminal situations involving mental health crises, homelessness, addiction, and interpersonal conflicts. However, these different types of calls merit very different responses, and relying on police as universal first responders might not be optimal. In particular, expanding the responsibilities of officers to include the mediation and de facto social work that is required while responding to certain types of noncriminal calls might strain their resources and impact their effectiveness at their core responsibilities. At worst, a lack of training in how to effectively address these situations might result in unnecessary arrests or violence. Appropriately responding to these types of calls is essential for fostering safer communities and reducing the burden on law enforcement.

One possible solution that has gained traction over the last several years is the addition of non-police first responders to cities' emergency-response portfolios (Irwin and Pearl, 2020; Krider et al., 2020). In their most typical form these teams consist of a non-uniformed mental health worker and a medic who are dispatched by emergency services in response to 911 calls. Crisis response teams are specially trained to address immediate crises, especially those caused by substance abuse, mental health issues, and nonviolent interpersonal conflict. Advocates argue that these challenges are distinct enough from regular police work that they merit an entirely different set of 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 for 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 the behavior of 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), the evidence on crisis response teams is sparse (Dee and Pyne, 2022, 2024).

This paper investigates the causal impact of sending crisis response teams to respond to nonviolent 911 calls—especially those 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 two complementary research designs to study the effect of both the initial expansion as well as of a marginal call after the program has reached scale, contributing to a comprehensive assessment of the effects and possible scope of crisis response programs.

At the beginning of our study period, in 2014, CAHOOTS operated from the middle of the morning to late at night in Eugene. Over the following four years, it expanded in five distinct steps to offer 24 hour coverage in Eugene as well as in the contiguous city of Springfield.The program originated in Eugene, Oregon, and gradually expanded in hours and geography to the nearby city of Springfield. Our first research design exploits this variation in a difference-in-differences design where we use the expansions to instrument for a CAHOOTS response to 911 calls, revealing the effect of CAHOOTS responses induced by the expansions.

Using detailed administrative 911 call data, we find that CAHOOTS responses that happened as a result of the expansions reduced the likelihood that a call resulted in arrest by 24 percentage points 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. However, we see no impacts on arrests related to criminal activity or outstanding warrants, suggesting that the 911 center is appropriately targeting crisis response teams towards low-risk calls. Consistent with this, we find that a CAHOOTS response substantially increases the likelihood that a medical unit from EMS arrives on-site, suggesting that one of the channels through which CAHOOTS improves outcomes is appropriately marshaling other city services.

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 conditional on these fixed effects; however, we find that these instrumented CAHOOTS responses have a small and usually statistically insignificant effect on the likelihood of arrest and on most other outcomes.

To understand the differing effects across designs, we turn to an analysis of CAHOOTS' substitution patterns with other emergency responders. 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. We view this as evidence that after the expansions, when CAHOOTS responds to about 8% of 911 calls, the program may have reached sufficient scale to be able to respond to the vast majority of calls

that would otherwise merit a police response.

CAHOOTS can respond to calls alone or alongside the police. To formalize this, we introduce an econometric framework that models our setting as having an unordered treatment with four possible levels: no response, CAHOOTS, police, and both. We show that under natural restrictions on response types, our IV analysis can reveal the composition of response types for each design, i.e. the share of instrumentally-induced responses where CAHOOTS acts as a substitute (responding instead of the police), a complement (responding alongside the police), or a service expansion (going to calls that otherwise would have received no response). Consistent with the smaller effects on arrests, the availability-induced responses are much less likely to be complements or substitutes than the expansion-induced responses, and much more likely to be service expansions.

To further unpack the extent to which substitution or complementarity is driving the reductions in arrest, we discuss conditions under which the "subLATEs" of each effect can be identified (Kline and Walters, 2016). Under the assumption that the subLATEs 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 using a simple over-identified 2SLS regression. After conducting a battery of statistical tests that suggest that this identification condition is reasonable in our context, we find that the reductions in arrests arise principally from complementarities between the police and CAHOOTS, rather than from a substitution effect.

Finally, we assess the effect of CAHOOTS on the likelihood of future calls from the same location, interpreting this as a measure of problem-solving. CAHOOTS responses reduce the likelihood of a follow-up call in the subsequent weeks, with the effect being larger and statistically significant for responses prompted by the introduction of CAHOOTS. These findings suggest that the reductions in arrests associated with CAHOOTS do not negatively impact crime or public safety.

This paper contributes to several areas of research. First, there has been much recent attention paid to the effects of the identity and behavior of emergency responders. The race and sex of officers affects stops, arrests, and use of force (Ba et al., 2021; Hoekstra and Sloan, 2022), and is particularly important for domestic violence cases (Miller and Segal, 2019). 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; Weisburst, 2019; Mello, 2019; Chalfin et al., 2022; Mello, 2024). 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, benefiting both the individuals involved and the officers responding to the incidents. The de-escalation of immediate crises can have significant benefits given recent work that highlights that high-profile police violence incidents lead to mistrust in the police and a rise in criminal behavior (Ang et al., 2023; Mikdash and Zaiour, 2024).

Another body of research, primarily descriptive research in medical journals, has directly examined crisis response teams, with much of the literature focusing on co-response models, where a police officer and a mental health professional (such as a psychiatric nurse or social worker) jointly respond to incidents. 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).

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 reports of the types of offenses targeted by the program (disorderly conduct, alcohol and drug use, and trespassing) by 34 percent through some combination of changed behavior and lower rates of reporting by STAR personnel. We complement this paper in several ways. First, we use detailed call data to move beyond a reduced-form analysis of the overall effect, and probe whether crisis responders should be best understood as a complement or substitute to the police. Second, we study the mechanisms through which crisis responder programs operate, such as bringing in other services such as EMS, or freeing up police officers to respond to other calls more quickly. Third, we can differentiate between changes in reporting and changes in behavior by studying the effect of a CAHOOTS response on a measure of problem-solving that is uncontaminated by reporting issues—future

calls from the same address. 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 marginal expansions to program capacity.

Finally, this is the first empirical paper to study the effect of CAHOOTS on 911 call outcomes. This contribution is particularly important given the outsized role that CAHOOTS 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.

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, CAHOOTS 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.[1] Crisis workers are not necessarily licensed professionals, but have at least two years of experience working in a non-traditional mental health setting, such as working with homeless outreach programs or crisis phone lines. Team members are unarmed and do not wear a uniform. Instead of police training, they complete specialized training, studying topics such as de-escalation, harm reduction, behavioral health, and substance abuse. Following classroom training, 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 on his/her own two-person team. The team's goal during each call is to address callers' immediate needs, rather than provide long-term solutions. This might include, for example, transporting callers experiencing a mental health crisis to either a health care facility or

[1]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.

a friends house, as the situation warrants. However, CAHOOTS also connects individuals with 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, the police non-emergency line, or the CAHOOTS non-emergency line. The calltaker's job is to talk to the caller to understand the situation and dispatch the appropriate responder. 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 natures are listed in Panel A of 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.

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

Our analysis relies on computer-aided dispatch (CAD) data from the Eugene and Springfield police departments. We link this with arrest and use of force records for each city. Data for Eugene and Springfield includes all calls between January 1, 2014 and December 14, 2021. The data includes both emergency and non-emergency calls. We focus on emergency 911 calls unless specifically mentioned otherwise. 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 or use of force.

For calls with an arrest, we can also see 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 being a threat to themselves or others(Gagnon et al. (2022)).

We merge call data with census tract information. For each census tract, we observe poverty and unemployment rates as well as demographic information.

2.4 Sample construction

Although our data includes both 911 calls and non-emergency police line calls, our analysis of call responses (i.e., whether CAHOOTS or the police are dispatched) and outcomes focuses solely on 911 calls. This decision is driven by several factors. First, our primary objective is to examine the substitutability and complementarity between CAHOOTS and police responses and how these dynamics influence call outcomes. Non-emergency calls, however, are less likely to result in a police dispatch. Second, after the introduction of CAHOOTS, many individuals began calling the non-emergency line specifically to request transportation. Unlike police officers, CAHOOTS teams can transport individuals to a variety of destinations based on their needs, such as hotels, urgent care facilities, physician offices, detox centers, community crisis centers, and housing programs. This means that the composition of the incoming non-emergency calls changed as a result of the CAHOOTS expansions, stymieing efforts to estimate difference-in-differences models at the call level. For 911 calls, however, we show below there was no change in call composition as a result of the expansions.

2.5 Summary statistics

Table 1 reports information on 911 calls in Eugene and Springfield from January 1, 2014 to December 14, 2021. Panel A reveals that CAHOOTS and the police are responsible for very different types of public safety emergencies. 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 also reveals differences in call outcomes. Calls receiving both CAHOOTS and police responses are substantially less likely to result in an arrest or use of force. 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. 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. CAHOOTS is usually on the scene 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.

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.[2] Table A2 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.[3] Each of the changes are permanent, with the number of calls slightly increasing

[2]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.

[3]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.

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:

$$Y_i = \beta \overbrace{D_i}^{\substack{\text{CAHOOTS} \\ \text{response}}} + \psi^1_{t(i)} + \gamma^1_{c(i)} + u_i \tag{1}$$

$$CAHOOTS_i = \alpha \underbrace{Z_i}_{\substack{\text{CAHOOTS} \\ \text{expansion}}} + \underbrace{\psi^0_{t(i)} + \gamma^0_{c(i)}}_{\substack{\text{Calendar month} \\ \text{\& city FEs}}} + \varepsilon_i \tag{2}$$

where Y_i is an outcome such as a recorded arrest as a result of call i, D_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-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 A2 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:

$$Y_{ij} = \beta D_{ij} + \psi^1_{t(i)j} + \gamma^1_{c(i)j} + \sum_{j' \neq j} \pi_{jj'} Z_{ij'} + u_{ij} \tag{3}$$

$$D_{ij} = \alpha_j Z_{ij} + \psi^0_{t(i)j} + \gamma^0_{c(i)j} + \sum_{j' \neq j} \alpha_{jj'} Z_{ij'} + \varepsilon_{ij} \tag{4}$$

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. 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 3 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, where the weights are proportional to the number of observations in the design multiplied by the residual variance in the instrument-predicted treatment.

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 expansion instrument. We find no statistically significant effects on these call natures, nor on predicted measures of a police response, arrest or an involuntary hold. We do find a small imbalance with respect to a predicted CAHOOTS response. However, in robustness analysis below, we show that controlling for these predicted call natures does not impact our estimated effects.

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 statistically significant relationships 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 significant confounding factors.

3.3 Estimated CAHOOTS effects on call response and arrests

Panel A of Table 3 reports reduced form estimates of the effect of an expansion on call outcomes. Before the expansions, CAHOOTS responded to about 0.5% of calls, which occurred either when the caller phoned early in the morning and waited until CAHOOTS started work for the day, or when CAHOOTS worked past their regular hours. 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 percentage points (F=178).

Panel B 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 39.6 percentage points relative to a control complier mean (CCM) of 96.6%, indicating that the expansion-induced CAHOOTS responses crowd out a considerable amount of police response. However, 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 gotten the police).

The remainder of Panel B estimates the effect on call outcomes. A CAHOOTS response dramatically reduces the likelihood of arrest, by 24 percentage points relative to a CCM of 32%. However, this reduction is almost entirely explained by involuntary holds (15 percentage points) and detoxification holds (5 percentage points), 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.

3.4 Estimated effects on access to medical services

Emergency responders are able to request additional emergency services if needed when they are responding to an emergency 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 fire or medical unit also responds to the call and on health outcomes.

Panel A of Table 5 shows estimates using our first research design that instruments for a CAHOOTS response using expansions in CAHOOTS services. Sending CAHOOTS increases the probability that a fire or EMS unit also responds to a call by 26 percentage points, relative to a control complier mean close to zero. These additional fire/EMS responses lead to a statistically significant 4.4 percentage point increase in the probability that a cardiac arrest is identified and a statistically significant 8.8 percentage point 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 effecive marshalling of other emergency services.

Table 5, Panel B shows that the 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 expansions research design. The descriptive correlations indicate that a CAHOOTS response is associated with a 1 percentage point reduction in fire/EMS responses, increases in suicidal persons, overdoses, individuals being dead on arrival, and a reduction in ER transports.

3.5 Estimated effects on future calls from the same address

Our analysis up until now has revealed that dispatching CAHOOTS has a profound effect on the call itself, with a CAHOOTS response substantially reducing the likelihood of an arrest. However, one possible concern is 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.

Ideally, one would like to study the effect of a CAHOOTS response on the subjects of 911 calls, to try to understand whether a CAHOOTS response made them more or less likely to be the subject of a future call or have any other interactions with the criminal justice system. In our context, however, this is infeasible, because the names of subjects are only collected when an arrest is made. Instead, we measure the need for future 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.

The red line in Panel A of Figure 4 reports the results using the expansions IV design. If anything, 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 33rd day after the focal call, after which they get smaller in magnitude and approach zero in the long term. 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 percentage points relative to a control complier mean of about 55% (Panel B).[4]

3.6 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 A4 shows the results, with Panel A replicating our preferred specification. Panel B controls for these tract characteristics, while 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 -0.226 to -0.241.

[4]To probe the robustness of our measure of future calls, Figure A4 Panel A measures future call 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 casese, we see declines in the 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.

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 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. 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($p = 0.85$; see Table 3).

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 3 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 A5), 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 making CAHOOTS an option relative to a counterfactual where the police are the only available emergency service.[5] 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

[5]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.

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 we call busyness. To control for potential confounding factors, we include detailed fixed effects at the city-calendar month-day of the week-hour of the day level. These comparisons focus, for instance, on 911 calls received in Springfield between 10 and 11 PM on Wednesdays in March 2021, comparing times with varying levels of CAHOOTS availability during that hour. To simplify visual representation and scaling we normalize busyness to be mean zero and with a standard deviation of one. Moreover, our main 2SLS specification binarizes the instrument to an indicator for having an above median availability when the focal call comes in.[6] We use a 2SLS specification, similar to Equation 1, in which a CAHOOTS response is instrumented by our measure of CAHOOTS availability:

$$
\begin{aligned}
Y_i &= \beta D_i + \gamma^1 Z_i^P + \eta^1_{h(i)} + \varepsilon_i \\
D_i &= \alpha \underbrace{Z_i^C}_{\substack{\text{High CAHOOTS} \\ \text{capacity}}} + \gamma^0 \underbrace{Z_i^P}_{\substack{\text{Police} \\ \text{busyness}}} + \eta^0_{h(i)} + \varsigma_i
\end{aligned}
\tag{5}
$$

where Y_i is an outcome such as a recorded arrest during call i, Z_i^C is an indicator for whether the call is in a time that CAHOOTS has high availability, Z_i^P is the number of open police cals, and $\eta_{h(i)}$ are 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 busyness on the likelihood of a CAHOOTS response using the continuous measure of CAHOOTS busyness. The relationship is negative and nearly monotonic; an additional standard deviation (SD) of busyness (approximately 3.5 additional calls) increases the likelihood of a CAHOOTS response by 2.1 percentage points. A natural concern is that busyness 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 specification in Equation 5.

[6] Working with a binary instrument makes the comparison to the difference-in-differences results from Section 3 more straightforward and is consistent with the conceptual econometric framework that we discuss in Section 5.1.

The estimates in Figure 2 show that calls that come at times with more capacity 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 characteristics and call natures. A CAHOOTS response is cross-sectionally associated with a 2 percentage point lower predicted arrest rate (mostly reflecting the difference between a call that is responded to at all versus not), 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 percentage points.

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 binary version of the availability instrument leads to a sharp increase of 3 percentage points in the likelihood of a CAHOOTS response with an F-statistic of 526. Unlike the expansions instrument, marginal increases in CAHOOTS are largely the result of responding to calls that otherwise would have not been answered, with high availability increasing the likelihood of a call response by 1.6 percentage points.

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 percentage points relative to a control complier mean of 45.9%, indicating that the availability-induced CAHOOTS responses crowd out a considerable amount of police response. However, 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. The effect of a CAHOOTS response on arrest is a 3.7 percentage point reduction that is not statistically significant. 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.

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.

4.4 Availability-induced dispatches and access to medical services

In this subsection, we measure the impact of a marginal CAHOOTS response on whether a fire or medical unit also responds to the call and on health outcomes. We saw above that CAHOOTS is much more likely to be providing an additional service for these marginal responses instead of serving as a complement or substitute to a police response.

Panel C of Table 5 shows estimates of the impact of CAHOOTS responses induced additional availability on these health outcomes. For these marginal expansions, a CAHOOTS response increases the probability of identifying a suicidal person by 4.7 percentage points, or by 82% relative to the control complier mean of 5.7%. However, a marginal CAHOOTS response does not affect the probability of an additional fire or EMS response, ER transport, or identifying an overdose, a cardiac arrest, or a deceased individual.

Mirroring the results in Panel B of Table 5, Panel D shows that the association between a CAHOOTS response and these health outcomes, controlling for the same set of controls used in our busyness design, looks quite different from the causal effect of marginal responses. We take this as another reminder of the importance of credible research design in this setting.

4.5 Availability-induced effects on calls from the same address

Next, we examine effects on future calls from the same address, interepreting these estimates as a potential measure of problem solving. The blue line in Figure 4 shows the results using the availability IV design. 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 whether there is a police response—smaller in magnitude than

the difference-in-differences estimates. In the longer term, the point estimates are close to zero and are statistically insignificant. While smaller in magnitude than the difference-in-differences estimates, the availability estimates again indicate that 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.

5 Estimating complementarity and substitution

5.1 An econometric framework for interpreting effects

The effect of a CAHOOTS response triggered by lower-than-usual busyness 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, reduces arrests by nearly four 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. To clarify these dynamics, we present a simple econometric framework.

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.[7] These restrictions can be formally stated as:

[7]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 busyness instrument as an additional control in robustness analysis below.

Assumption 1 (IA monotonicity)**.**

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

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

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

The allowable response types are summarized in Table 6. 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.

Although the specific treatment effects for each of the complier types are not identified without further assumptions, the shares of each of these response types in the overall population, $\pi_{xy} \equiv P[D(0){=}x, D(1){=}y]$, are identified from changes in observed treatments:

$$\begin{aligned}
\pi_{pb} &= E\big[\mathbb{1}[D_i{=}b] \mid Z_i{=}1\big] - E\big[\mathbb{1}[D_i{=}b] \mid Z_i{=}0\big] \\
\pi_{pc} &= E\big[\mathbb{1}[D_i{\in}\{n, c\}] \mid Z_i{=}1\big] - E\big[\mathbb{1}[D_i{\in}\{n, c\}] \mid Z_i{=}0\big] \\
\pi_{nc} &= E\big[\mathbb{1}[D_i{\in}\{c, p, b\}] \mid Z_i{=}1\big] - E\big[\mathbb{1}[D_i{\in}\{c, p, b\}] \mid Z_i{=}0\big]
\end{aligned} \tag{6}$$

As a first step towards characterizing possible differences in treatment between the busyness and expansion designs, Figure 5 reports the shares of compliers from each response type, $s_{xy} = \pi_{xy}/(\pi_{nc} + \pi_{pc} + \pi_{pb})$. 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. One testable implication of Assumption 2 is that each of the estimated π's should be positive; Figure 5 shows that this condition is satisfied.

We view these results as suggesting that there is a relatively small 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—for example by adding additional vans during the day—would likely match this pattern.

5.2 What do LATE estimates recover?

These assumptions allow us to more sharply characterize what we can learn from IV regressions in this environment. Abstracting from controls, we focus on the Wald estimator of the effect of a CAHOOTS response ($\mathbb{1}[D \in \{c, b\}]$) that we examined in previous sections, and show that the estimand is composed of the weighted average of the effects for the thee groups: (1) the *complementarity* of CAHOOTS and the police in working together (Δ_{pb}) ; (2) a *substitution* effect from replacing the police with CAHOOTS (Δ_{pc}); and (3) the impact of *increasing services* to calls the police would not respond otherwise (Δ_{nc}):

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

$$+ \underbrace{\frac{\pi_{pc}}{\pi_{nc}+\pi_{pc}+\pi_{pb}}}_{\equiv s_{pc}} \overbrace{E[Y(c) - Y(p)|D_i(1)=c, D_i(0)=p]}^{\equiv \Delta_{pc}} + \underbrace{\frac{\pi_{nc}}{\pi_{nc}+\pi_{pc}+\pi_{pb}}}_{\equiv s_{nc}} \overbrace{E[Y(c) - Y(n)|D_i(1)=c, D_i(0)=n]}^{\equiv \Delta_{nc}}$$

Substituting from police to CAHOOTS

CAHOOTS response to a call the police would not (7)

Equation 7 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 the composition of complier types s_{pc}, s_{pb}, and s_{nc}.

The estimand in Equation 7 recovers a well-defined causal parameter. However, for some outcomes, like arrests, the last term will mechanically be zero, and by rescaling the reduced form, we can recover the LATE of a weighted average of Δpc and Δ_{pb} under the same identifying assumptions:

$$\frac{E[Y|Z=1] - E[Y|Z=0]}{E[1[D \neq p]|Z=1] - E[1[D \neq p]|Z=0]} = \underbrace{\frac{s_{pc}}{s_{pc}+s_{pb}} \cdot \Delta_{pc} + \frac{s_{pb}}{s_{pc}+s_{pb}} \cdot \Delta_{pb}}_{\text{Rescaled effect}} \tag{8}$$

Panel A of Table 7 reports estimates of the rescaled effect from Equation 8 and reproduces the estimates of Equation 7 from Table 3 for comparison. In the expansion design, the rescaled and non-rescaled estimates are very similar because the instrument has a negligible level of $n \to c$ compliers (s_{nc}). However, in the availability design, there is a meaningful share of $n \to c$ compliers and therefore the magnitude of the effect doubles when using Equation 8 rather than Equation 7. Nonetheless, the rescaled effects remain quite different for the availability and expansion designs (-0.088 and -0.241, respectively, $p = 0.14$). To explore whether this is caused by differences in the degree of substitution and complementarity across designs, or differences in subLATEs, we turn in the next section to directly estimating the subLATEs.

5.3 Estimating subLATEs

One way to view our setting is that with unrestricted heterogeneity in treatment effects, the problem is underidentified. In this section, we show how assumptions on the degree of allowable heterogeneity enable identification of the subLATEs.

There are five expansions used in the expansion design. The Wald estimate for each expansion e identifies the expansion-specific treatment effect:

$$s^e_{pc}\Delta^e_{pc} + s^e_{pb}\Delta^e_{pb} + s^e_{nc}\Delta^e_{nc}.$$

This results in a system of five equations with 15 unknowns. However, if we assume that treatment effects are constant across expansions (i.e., $\Delta^e_{pc}=\Delta_{pc}$ for all e), the system becomes an over-identified linear system with five equations and three unknown parameters: Δ_{pc}, Δ_{pb}, and Δ_{nc}. To improve efficiency, we can further impose $\Delta_{nc}=0$, as we know from our institutional context that an arrest cannot take place without the arrival of the police.[8]

Panel B of Table 7 reports the subLATEs, which we construct using the estimates from an overidentified 2SLS specification that instruments for $\mathbb{1}[D_i=c]$, $\mathbb{1}[D_i=p]$, and $\mathbb{1}[D_i=b]$ with the five post-expansion indicators. Column (2) reveals that the reductions in arrests arise mostly from $p \to b$ calls rather than $p \to c$ calls, although we cannot reject equality of the coefficients. The model also estimates an $n \to c$ effect of 0.08 (SE=0.33), which—since we know that the $n \to c$ effect has to be zero—we view as an implicit validation of the model. More directly, in Panel C we report the results of overidentification tests and find that we do not reject the null of homogenous treatment effects across the five expansion

[8] This constraint can be interpreted as a shape restriction motivated by our empirical setting (Chetverikov et al., 2018).

instruments.

In column (3) we directly impose Δ_{nc}=0 to improve power. The estimates reveal that complementing the police with CAHOOTS reduces arrests by a statistically significant 0.36, while replacing the police with CAHOOTS reduces arrests by only 0.10. While these effects are not directly comparable in the sense that they are valid for different response types, they nonetheless indicate that the benefits of CAHOOTS in reducing arrests comes largely from them acting as a complement to the police, rather than a substitute.

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 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, and in the case of the rescaled effect we reject equality with the empirical estimates at the 10% level. We conclude 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 Cost trade-offs between CAHOOTS and police

Our results indicate that approximately 8 percent of 911 calls are suitable for a CAHOOTS response. Of these, CAHOOTS substitutes for police in about 4% of cases and complements police in another 4%.

CAHOOTS offers significant cost savings when it replaces police. In 2024, the hourly wage for a new police officer in Eugene was $36.35, increasing to $46.34 by the sixth step of the pay scale.[9] In contrast, CAHOOTS responders earn a base wage of $18 per hour (Eugene Register Guard, 2024). Therefore, the marginal cost of deploying a CAHOOTS team is at most half that of sending two police officers. In fact, the total cost of a CAHOOTS response—including salaries, taxes, training, equipment, and insurance—is comparable to

[9] https://www.eugene-or.gov/DocumentCenter/View/2387/City-of-EugeneSalary-Schedule

the labor cost of two new police officers ($72.33 per hour).[10]

For the remaining 4% of cases where CAHOOTS complements police, there are no direct cost savings. However, indirect savings—such as fewer arrests—may offset the additional cost. On average, a 911 call ending in an arrest takes about 3 hours, compared to 38 minutes for calls without arrests. That said, the impact of expanding CAHOOTS may be limited, as we find only a small, statistically insignificant reduction in arrests for these calls.

7 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. 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.

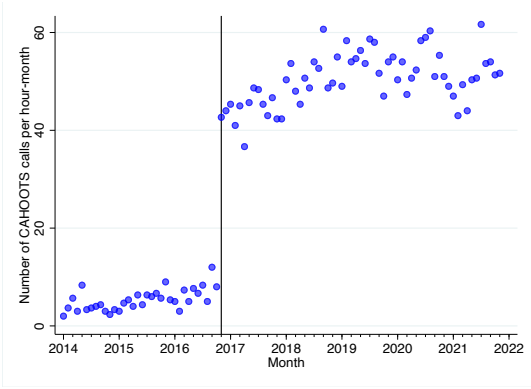
However, CAHOOTS' ability to serve as a substitute for the police is limited. We find that in the initial expansions that made CAHOOTS suddenly available in new times of day or geographic areas, about half of CAHOOTS responses were substitutes for the police (in the other half, CAHOOTS responded alongside the police). However, after the initial expansions, when CAHOOTS responds to about 8% 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.

[10] In FY 2022, the city of Springfield paid $459,500 for CAHOOTS services, with 6,353 responses in 2021. This translates to a cost of $72.33 per response. Source: https://springfield-or.gov/wp-content/uploads/2022/06/FY22-CITY-Adopted-Budget_6.15.22.pdf

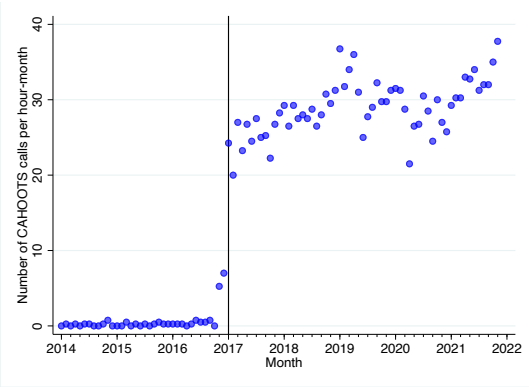
Figures

Figure 1: CAHOOTS expansions in Eugene and Springfield

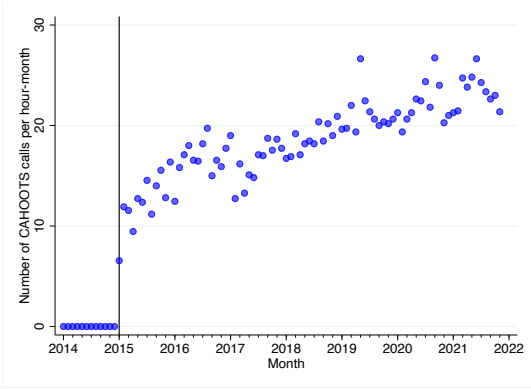
(a) Eugene 7-10 a.m.



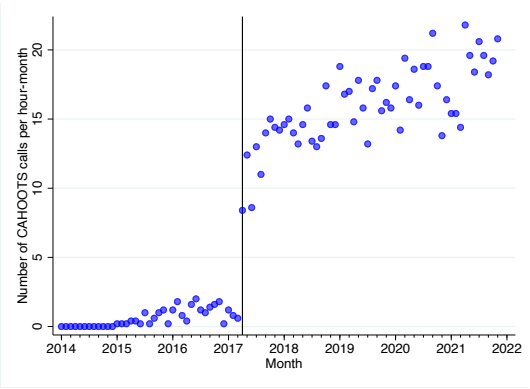
(b) Eugene 3-7 a.m.



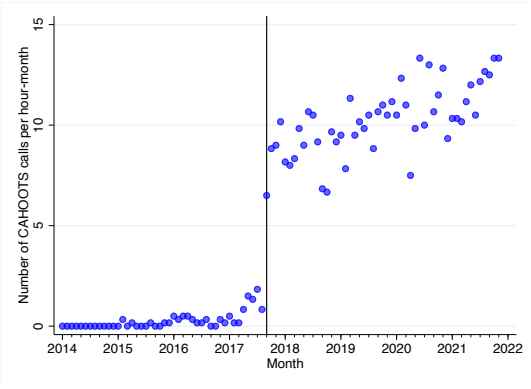
(c) Springfield 12 p.m. to 11 p.m.



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

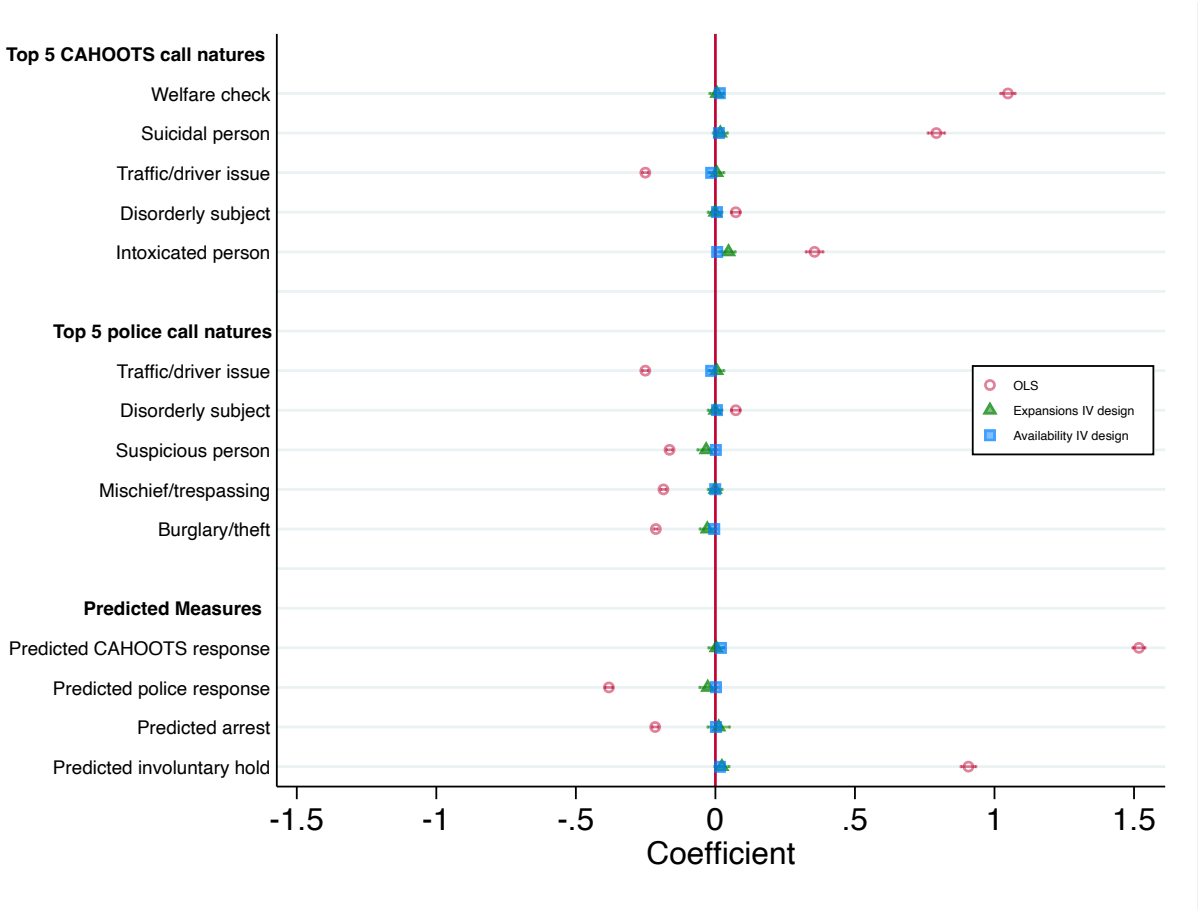


(e) Springfield 3-10 a.m.



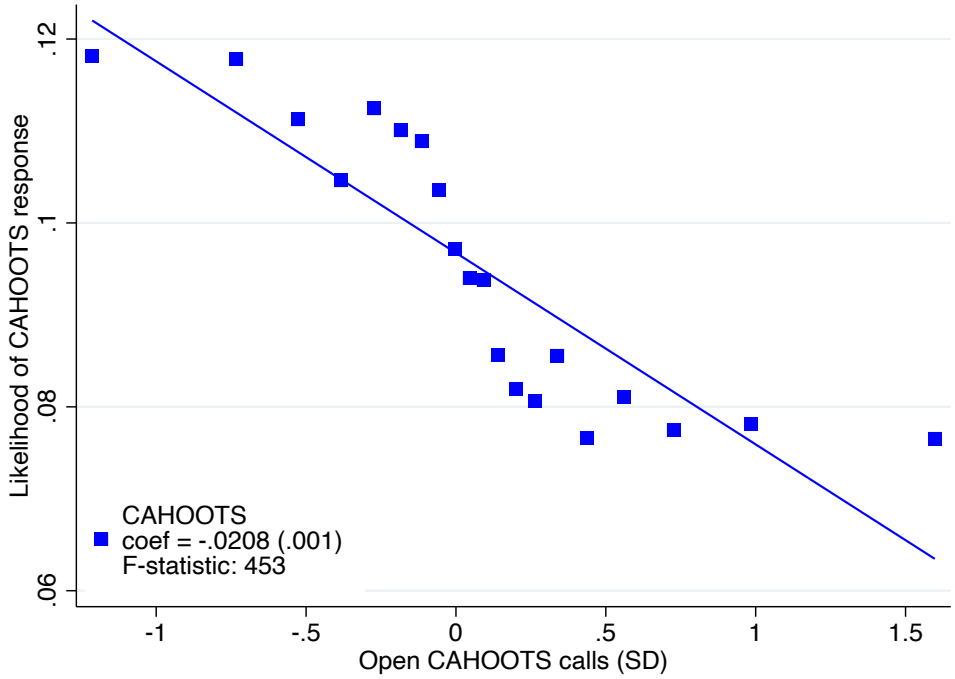
Notes: This figure shows the changes in the number of calls to which CAHOOTS is dispatched as services are expanded (Eugene) or introduced (Springfield). Panels A and B report expansions in Eugene and Panels C, D, and E expansions in Springfield. Each point in the figures reports the number of calls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

Figure 2: Balance checks of CAHOOTS expansions and busyness instrument



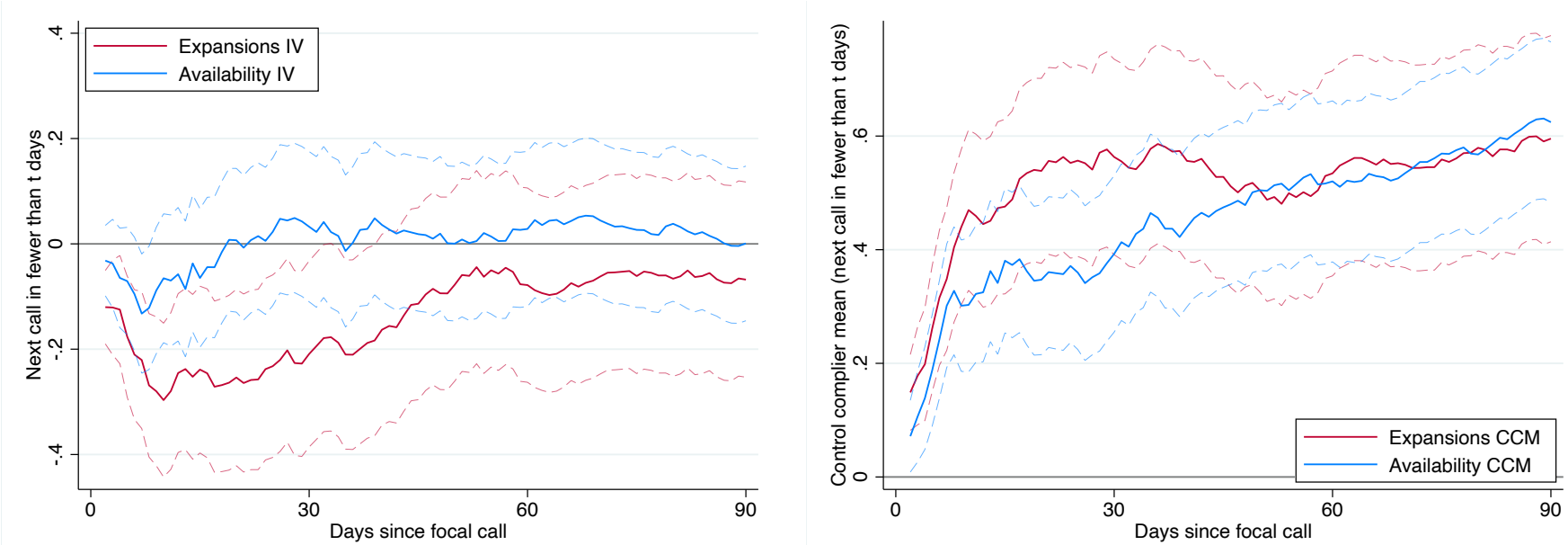
Notes: The figure reports estimates of the association between CAHOOTS response and call natures as well as predicted measures based on natures and the geographic characteristics of the call location (red hollow circle markers). It also includes the association between CAHOOTS expansions and call characteristics (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: First stage of CAHOOTS busyness instrument



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

Figure 4: Effect of CAHOOTS response on future call from same address

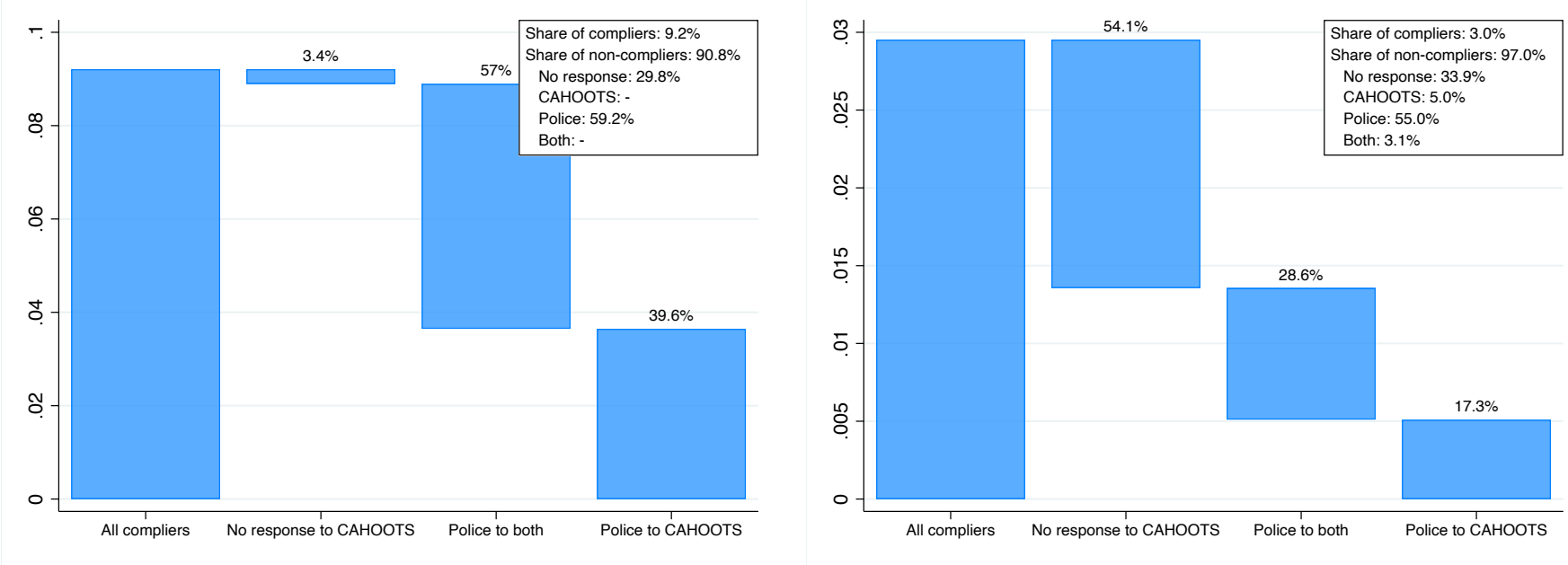


A. Effect on future call within t days

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.

Figure 5: 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.

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 average call characteristics 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, 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 using expansions design

	Call responses			Call outcomes			
	CAHOOTS	Police	Any response	Arrest	Invol. hold	Detox	Warrant
	(1)	(2)	(3)	(4)	(5)	(6)	(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 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Effects of CAHOOTS 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 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: 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 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: 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 $\pi_{xy} \equiv Pr[D(0)=x, D(1)=y]$. 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 7: Complier group-specific treatment effects on arrest

	Instruments		
	Availability	Diff-in-diff	
	(1)	(2)	(3)
Panel A: Treatment effect estimates			
CAHOOTS effect ($s_{pc}\Delta_{pc} + s_{pb}\Delta_{pb} + s_{nc}\Delta_{nc}$)	-0.0395	-0.240	-0.240
	(0.0406)	(0.0614)	(0.0614)
Rescaled effect ($\frac{s_{pc}}{s_{pc}+s_{pb}}\Delta_{pc} + \frac{s_{pb}}{s_{pc}+s_{pb}}\Delta_{pb}$)	-0.0857	-0.241	-0.241
	(0.0861)	(0.0616)	(0.0616)
Panel B: subLATEs			
Police → both		-0.405	-0.357
		(0.274)	(0.165)
Police → CAHOOTS		-0.031	-0.100
		(0.394)	(0.224)
None → CAHOOTS		0.080	
		(0.334)	
$\Delta_{pb} = \Delta_{pc}$ (p-value)		0.569	0.486
Panel C: Goodness-of-fit tests			
J test of overidentification			
χ^2 statistic		0.351	0.409
p-value		0.839	0.938
Predicted availability design LATEs using subLATEs			
CAHOOTS effect ($s_{pc}^{busy}\Delta_{pc} + s_{pb}^{busy}\Delta_{pb} + s_{nc}^{busy}\Delta_{nc}$)		-0.078	-0.120
		(0.180)	(0.029)
p-value of equality		0.834	0.103
Rescaled effect ($\frac{s_{pc}^{busy}}{s_{pc}^{busy}+s_{pb}^{busy}}\Delta_{pc} + \frac{s_{pb}^{busy}}{s_{pc}^{busy}+s_{pb}^{busy}}\Delta_{pb}$)		-0.263	-0.260
		(0.065)	(0.063)
p-value of equality		0.093	0.098
Predicted diff-in-diff design LATEs using subLATEs			
CAHOOTS effect ($s_{pc}^{DD}\Delta_{pc} + s_{pb}^{DD}\Delta_{pb} + s_{nc}^{DD}\Delta_{nc}$)		-0.240	-0.243
		(0.061)	(0.059)
p-value of equality		1.000	0.862
Rescaled effect ($\frac{s_{pc}^{DD}}{s_{pc}^{DD}+s_{pb}^{DD}}\Delta_{pc} + \frac{s_{pb}^{DD}}{s_{pc}^{DD}+s_{pb}^{DD}}\Delta_{pb}$)		-0.252	-0.251
		(0.061)	(0.061)
p-value of equality		0.586	0.587

Notes: This table presents estimates of complier-specific treatment effects on arrest. The first row of Panel A reports our baseline estimate, which can be interpreted as a linear combination of the effect of a CAHOOTS response rather than the police (Δ_{pc}), the effect of both the police and CAHOOTS responding rather than the police alone (Δ_{pb}), and the effect of a CAHOOTS response rather than no response (Δ_{nc}). The second row presents the rescaled effect of moving from a police-only response. Panel B reports subLATEs estimated under the assumption of compliance-group specific constant treatment effects across expansions. Panel C reports goodness-of-fit tests, beginning with the standard J test of treatment effect homogeneity across instruments. We also use the subLATEs to construct reweighted treatment effects using (respectively) the weights from the availability design and the expansion design .

References

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. Bruhn, and E. Derenoncourt (2023). *Shots fired: Crime and community engagement with law enforcement after high-profile acts of police violence*. Benjamin H. Griswold III, Class of 1933, Center for Economic Policy Studies ....

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., 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.

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.

Chetverikov, D., A. Santos, and A. M. Shaikh (2018). The econometrics of shape restrictions. *Annual Review of Economics 10*(1), 31–63.

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.

Dee, T. S. and J. Pyne (2024). Mental-health first responders reduce involuntary psychiatric detentions. Unpublished manuscript.

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-garners73225506007/.

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.

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.

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.

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.

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.

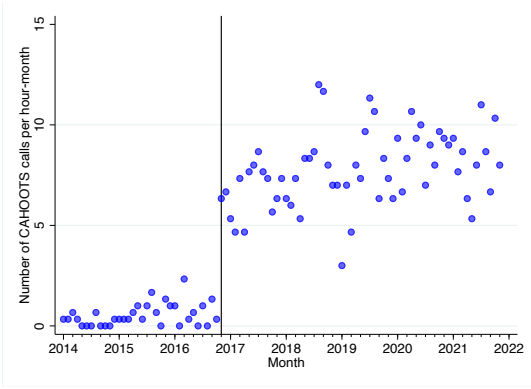
Moreno-Medina, J., A. Ouss, P. Bayer, and B. A. Ba (2023). Officer-involved: The media language of police killings. Technical report, Accepted *Quarterly Journal of Economics*. Available at https://static1.squarespace.com/static/5ec5aa4736bba52179f2ff85/t/65dcf3227d4dbf14c32d1f3a/1708978979966/police_language.pdf.

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.

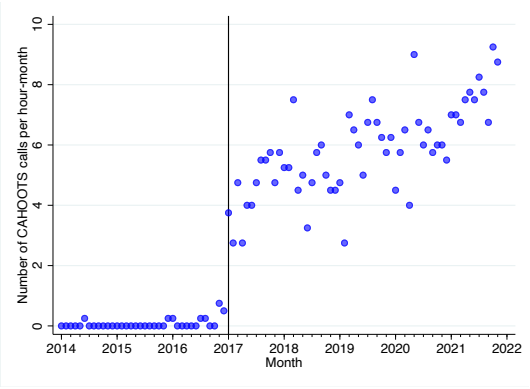
A1 Appendix Figures

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

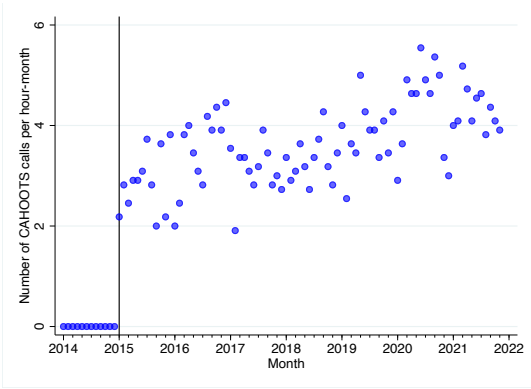
(a) Eugene 7-10 a.m.



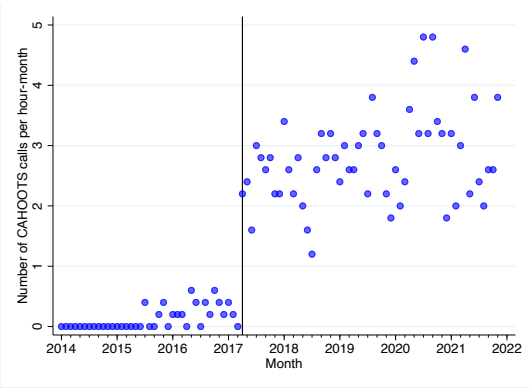
(b) Eugene 3-7 a.m.



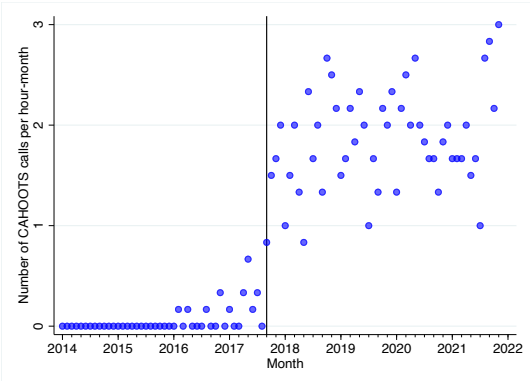
(c) Springfield 12 p.m. to 11 p.m.



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



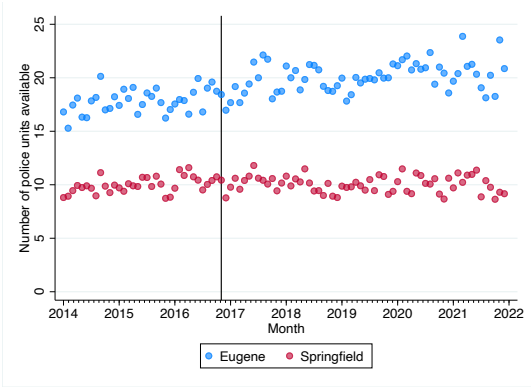
(e) Springfield 3-10 a.m.



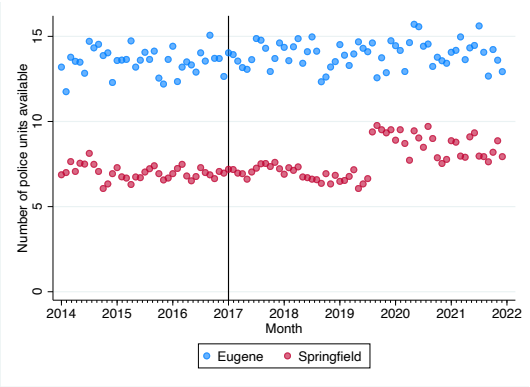
Notes: This figure shows the changes in the number of calls to which CAHOOTS is dispatched as services are expanded (Eugene) or introduced (Springfield). Panels A and B report expansions in Eugene and Panels C, D, and E expansions in Springfield. Each point in the figures reports the number of 911 calls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

Figure A2: Police units available during design hours

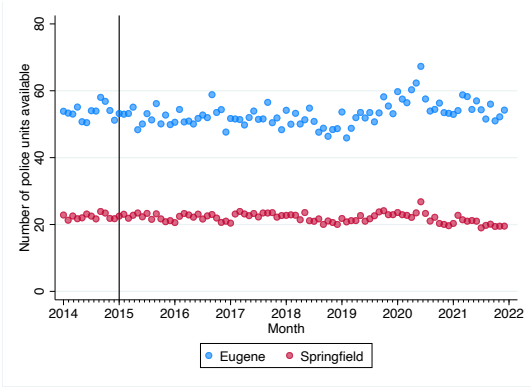
(a) 7-10 a.m.



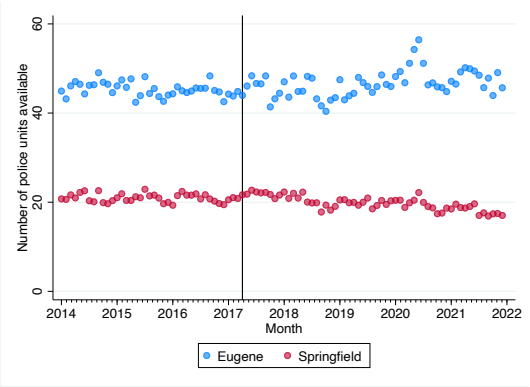
(b) 3-7 a.m.



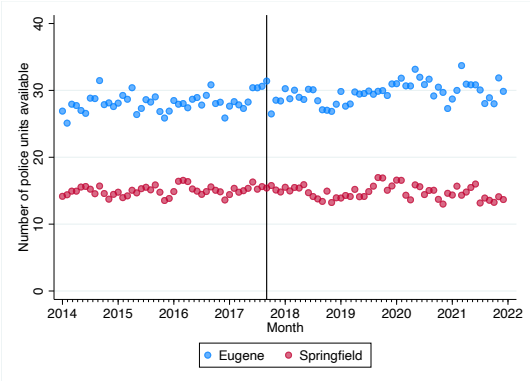
(c) 12 p.m. to 11 p.m.



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

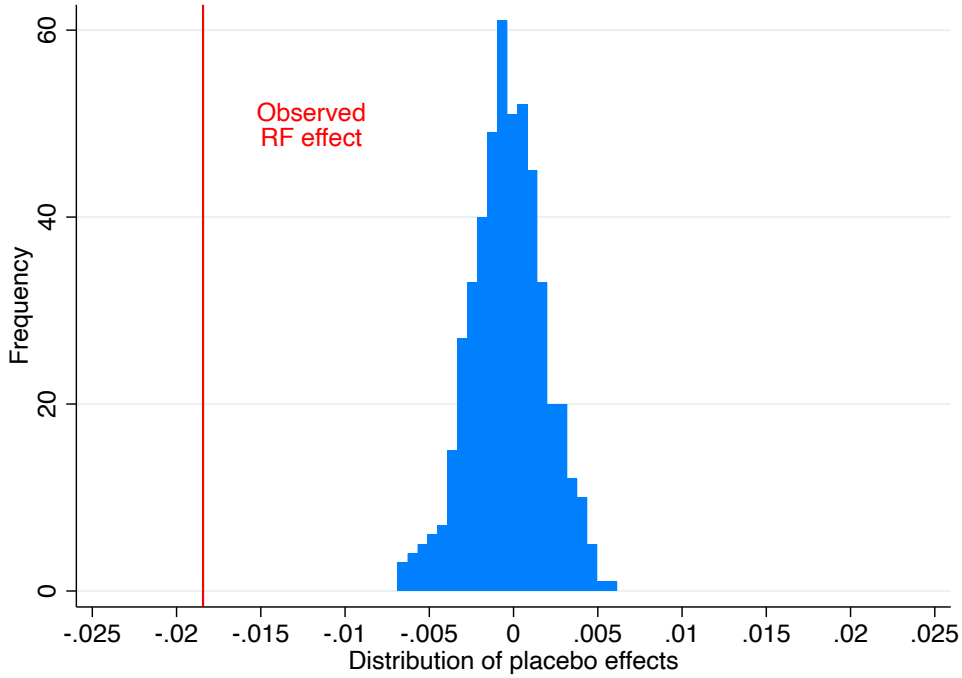


(e) 3-10 a.m.



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

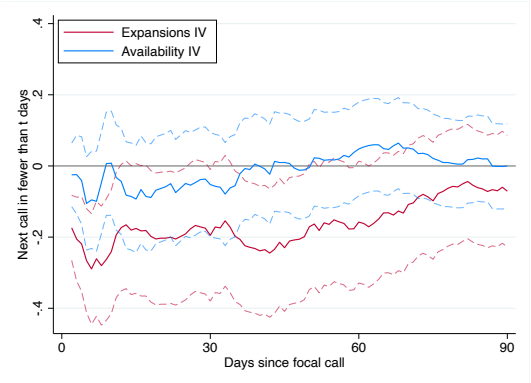
Figure A3: Distribution of placebo expansion times



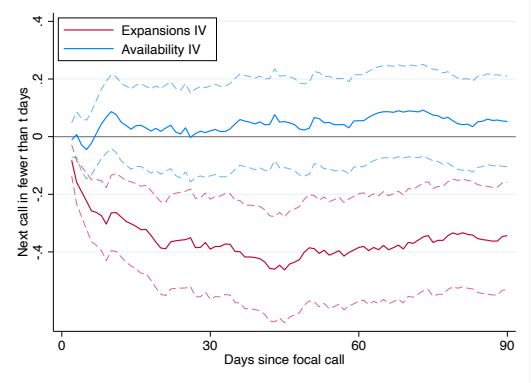
Notes: This figure displays the distribution of estimated effects for 500 randomly chosen event times. These effects are estimated using the same specification as in Equation 3, focusing on the reduced-form relationship. The placebo event time effects are small, centered around zero, and do not overlap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

Figure A4: Robustness for analysis of effect of CAHOOTS response on future calls from area in fewer than t days

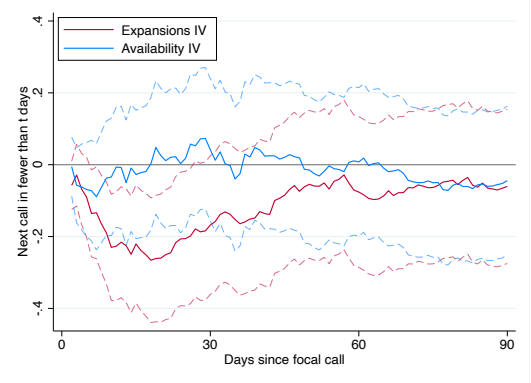
(a) Within 50 meters of focal call



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



(c) Excluding high-traffic and public places



Notes: Panel A shows the effect of 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.

A2 Appendix Tables

Table A1: Call characteristics by response

	(1) Count	(2) No response	(3) CAHOOTS only	(4) Police only	(5) Both
All calls	270,728	0.334	0.054	0.579	0.033
Welfare check	16,206	0.217	0.354	0.369	0.060
Suicidal person	6,559	0.131	0.245	0.370	0.254
Disorderly subject	20,070	0.209	0.044	0.683	0.064
Dispute	43,042	0.103	0.002	0.864	0.032
Traffic/driver issue	31,779	0.490	0.023	0.478	0.009
Intoxicated person	1,399	0.204	0.415	0.313	0.067
Transportation	996	0.225	0.568	0.151	0.056
Suspicious person	17,097	0.243	0.014	0.721	0.023
Overdose	4,018	0.423	0.031	0.471	0.075
Mischief/trespassing	13,251	0.324	0.007	0.649	0.020
Mental issue	2,031	0.113	0.109	0.715	0.063
Subject down	1,391	0.410	0.171	0.373	0.045
Assault	5,812	0.116	0.003	0.835	0.046
Cardiac arrest	2,707	0.478	0.004	0.446	0.071
Burglary/theft	11,269	0.369	0.001	0.622	0.008
Harassment	3,884	0.246	0.003	0.736	0.015
Incomplete call	13,092	0.724	0.002	0.271	0.003
Warrant	1,451	0.001	0.000	0.971	0.028
Armed subject	1,409	0.103	0.001	0.869	0.026
Drunk driver	9,735	0.624	0.000	0.374	0.002
Fight	1,688	0.094	0.000	0.892	0.014
Robbery	868	0.094	0.001	0.879	0.025
Shots fired	3,417	0.353	0.001	0.644	0.003

Notes: This table shows average call characteristics (call nature) 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

Table A2: 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 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.

Table A4: 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 A5: 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 approach, all 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 A6: 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 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.00$