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

September 2024

Abstract

This paper studies the use of mobile crisis teams—a non-uniformed pair consisting of a mental health worker and a medic—as either a substitute or complement to traditional police units responding to 911 calls involving mental illness, homelessness, addiction, and other non-dangerous situations. We evaluate the longest-running mobile crisis response program in the U.S., Crisis Assistance Helping Out on the Streets (CAHOOTS) in Eugene, Oregon. We find that a series of expansions of CAHOOTS into new areas and times reduced the likelihood that a 911 call resulted in an arrest, and had the long-term benefit of reducing the number of follow-up 911 calls from the same address. However, CAHOOTS often responds to the same calls as the police, acting as complement rather than a substitute. The reduction in arrests likely reflects CAHOOTS' role in de-escalating tense situations and resolving incidents without coercive measures. After the initial expansions, additional CAHOOTS capacity is used mostly for calls that would otherwise go unanswered. We conclude that crisis response teams play an important role as a complement to the police, rather than only acting as substitutes. Further expansions are likely to have only limited effects on reducing police responses.

*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 Felipe Goncalves, Jens Ludwig, and Grant McDermott, and seminar participants at multiple universities for their helpful comments. Diti Jain and Everett Stamm provided outstanding research assistance.

Police officers respond to a wide variety of 911 calls, from urgent reports about violent crimes in progress all the way to noncriminal situations involving mental illness, homelessness, addiction, and interpersonal disputes. However, some activists and advocates have argued that the police are ill-suited to respond to many of these noncriminal situations, and that in some situations their involvement may be counterproductive or lead to unnecessary violence. If the city had other means to respond to these calls, furthermore, police resources could be diverted to improve response times or respond to other calls.

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. In their most typical form, these teams consist of a non-uniformed mental health worker and a medic, and are dispatched by emergency services in response to 911 calls. These crisis response teams are specially trained to address immediate crises, especially those brought about by substance abuse, mental health issues, and non-violent interpersonal conflict. Advocates argue that these challenges are distinct enough from regular police work that they merit an entirely different set of responders.

Despite the large public demand for alternatives and complements to the police (Ba et al., 2024), there is little evidence of the causal effects of crisis responders on the outcome of 911 calls. Are they a substitute for the police, or do they respond to calls they would not have been dispatched to? Do they successfully de-escalate immediate crises and reduce arrests and use of force? And what are the longer-term effects of these teams on public safety? While there is extensive literature on the behavior of police officers (e.g. Anwar and Fang, 2006; Goncalves and Mello, 2021; Chalfin et al., 2022; Feigenberg and Miller, 2022; Moreno-Medina et al., 2023; Chalfin and Gonçalves, 2023, among others), the evidence on crisis response teams is much sparser (Dee and Pyne, 2022).

This paper investigates the causal impact of sending mental health workers and medics instead of police in response to 911 calls involving non-dangerous, non-criminal episodes. Our context is the longest-running mobile crisis response program in the United States, Crisis Assistance Helping Out on the Streets (CAHOOTS). The program originated in Eugene, Oregon and gradually expanded in hours and geography to the nearby city of Springfield. We exploit this variation in a difference-in-differences design where we use the expansions to instrument for a CAHOOTS response to 911 calls, allowing us to estimate causal effects on call outcomes like arrest and use of force, as well as the extent to which CAHOOTS acts as a substitute or complement to traditional police officers.

Using detailed administrative 911 call data, we find that the rollout of CAHOOTS reduced

the likelihood that a call resulted in arrest by 20 percentage points relative to a control complier mean of 29%, while having no impact on use of force. In this intial phase, CAHOOTS acted as both a subsitute and a complement to the police: of the 8% of calls that CAHOOTS responded to after the expansions, approximately half of them also had a police response. The other half, which CAHOOTS responded to alone, would have had a police response in the absence of CAHOOTS availability.

We also study the effect of further expanding CAHOOTS availability. Focusing on calls that occurred after the initial expansions, we instrument for the type of response—police or CAHOOTS—with the number of outstanding calls of each type at the time of the focal call. The response patterns are quite different in this case, with the majority of marginal CAHOOTS responses being for calls that otherwise would have received no police response. Consistent with this more limited extent to which CAHOOTS is either a subsitute or a complement to the police, we find same-signed but smaller and statistically insignificant effects on arrests and use of force.

Finally, we estimate the effect of the two types of emergency response on frequency and timing of future calls from the same address, which we interpret as a measure of whether the response successfully resolved the problem. A CAHOOTS response reduces the likelihood of a call in the next several weeks, although this effect is larger and only statistically significant for the responses induced by the initial expansions. A police response, in contrast, slightly increases the likelihood of a follow-up call.

This paper contributes to several areas of research. First, there has been much recent attention paid to the effects of changes to 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; Evans and Owens, 2007; Chalfin and McCrary, 2018; Mello, 2019; Chalfin et al., 2022). While increasing the number of police officers has been shown to reduce crime, recent work argued it could 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 not well-equipped to manage. Importantly, many of these im-

provements 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.

Another smaller literature has more directly studied crisis response teams. Dee and Pyne (2022) examine Denver's Support Team Assisted Response (STAR) program and find reductions in reported crime in precincts where STAR operates. Specifically, difference-in-difference estimates at the precinct-month show that the program decreased reports of the types of offenses targeted by STAR (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-level data to move beyond a reduced-form analysis of the overall effect of a crisis responder program, 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 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 an important measure of recidivism, future calls from the same address.

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 have piloted their own versions. As these programs continue to grow across the country, evidence on their effectiveness and interaction with existing emergency response services is key to their success.

2 Background and data

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 (nurse, paramedic, or EMT) and a mental health crisis worker—in a Sprinter-style van loaded with emergency supplies. Team

members are unarmed and do not wear a uniform or receive police training. Instead, they complete 500 hours of specialized training, studying topics such as de-escalation, behavioral health, and substance abuse. Their goal during each call is to address callers' immediate needs, rather than provide long-term solutions.

Table 1 reports information on 911 calls in Eugene and Springfield for the eight years from the beginning of 2014.[1] Restricting attention to the calls at a time and location where CAHOOTS was operating, the table reveals that CAHOOTS is responsible for very different types of public safety emergencies than the police.[2] CAHOOTS disproportionately responds to calls for a wellness check and suicidal persons, but is never assigned to serious or potentially violent situations such as fights, armed persons, or drunk drivers. Since sometimes multiple units are dispatched to a call, we also observe that there are some types of emergencies where both CAHOOTS and the police are often assigned, such as disorderly subjects and interpersonal disputes.

These differences in types of calls reflect the differing specialties and training of CAHOOTS and the police. For example, while police officers are supposed to transport individuals only to jail or a hospital, CAHOOTS teams will also transport callers to other locations according to their needs, including hotels, urgent care, physician's offices, detox centers, community crisis centers, and housing programs. In addition to connecting individuals to resources in the community, CAHOOTS provides a range of on-scene services. These include 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 1 also reveals differences in call outcomes. CAHOOTS-only calls (column (2)) take longer to be dispatched and longer for the first unit to arrive than calls to which the police are also sent (columns (3) and (4)); for example the average response takes 39 minutes for a CAHOOTS-only call and 19 minutes for a police-only call. This reflects that there is usually only one CAHOOTS van available, as well as the fact that the CAHOOTS calls are often less urgent. CAHOOTS calls are also 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, 14.2% of police-only calls and 14.1% of CAHOOTS-and-police calls result in an arrest. Similarly, CAHOOTS-only calls never result in the use of force, but a little less than 1% of calls involving the police do.

[1] Data for Eugene runs until December 14, 2021, and for Springfield until July 4, 2022.

[2] Table A4 contains summary statistics for the slightly larger total sample, which includes some calls in times or areas where CAHOOTS was not dispatchable.

At the beginning of our study period, CAHOOTS was available only in Eugene between 11 a.m. and 3 a.m. Over the next several years, CAHOOTS expanded its hours and geographic area of operations. Table A2 contains details of each of the five separate expansions that occurred over the four years from 2014, first to the neighboring town of Springfield, and then in several area-specific steps to 24 hour service. We begin by using these expansions in service to estimate the effect of the initial rollout of CAHOOTS on call outcomes.

3 Effects of the CAHOOTS rollouts

Our first empirical strategy exploits the five CAHOOTS expansions that took place during our study period. Each expansion made CAHOOTS vans suddenly available for additional hours on a specific date in a specific area, either Eugene or Springfield.[3] Before each expansion, the number of calls that were dispatched to CAHOOTS was low but non-zero, and were the result of either vans working past the end of their shift to address additional calls, or dispatchers waiting for CAHOOTS vans to begin their shifts so they could assign a call to them.

Figure 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 Eugene policy change 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. Each of the changes are permanent, with the number of calls slightly increasing over the subsequent years.

These graphs motivate our first empirical strategy, which uses the expansions as instrumental variables for CAHOOTS dispatch. We focus on 911 calls, both because they are less likely to be calls specifically requesting CAHOOTS and for consistency with our second identification strategy described in Section 4.[4] Each design $d(j)$ consists of the hours of the day affected by the j^{th} policy change, whether or not they are in the city where the policy change occurred. We stack designs and use each policy change as an IV. Since some of the designs overlap—for example, Table A2 reveals that the first Eugene expansion and the second Springfield

[3] 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.

[4] Figure A1 shows the number of 911 calls per hour-month for each policy change. The jumps are sharp but somewhat noisier, reflecting the lower volume of 911 calls.

expansion both included 9-10 a.m.—we additionally control for the other policy changes. More specifically, we estimate the following 2SLS specification

$$Y_{ij} = \beta D_{ij} + \phi_{d(j)t} + \theta_{d(j)c(i)} + \sum_{j' \neq j} \pi_{jj'} Z_{ij'} + u_{it} \tag{1}$$

$$D_{ij} = \alpha_j Z_{ij} + \psi_{d(j)t} + \gamma_{d(j)c(i)} + \sum_{j' \neq j} \alpha_{jj'} Z_{ij'} + \varepsilon_{it} \tag{2}$$

where Y_{ij} is an outcome such as a recorded use of force on call i, D_{ij} is an indicator for a CAHOOTS response, and Z_{ij} is an indicator for the call coming in a time and geography following the j^{th} expansion. t indexes months and $c(i)$ the two cities in our sample, so we control for design-month effects $\psi_{d(j)t}$ and design-city effects $\gamma_{d(j)c(i)}$, analogous to time and location fixed effects in a traditional diff-in-diff. Some observations appear twice in this regression because of the design overlap, so we cluster this regression at the call level.

Panel A of Table 2 shows how a CAHOOTS response changes the treatment received by the caller. As suggested by Figure 1, the policy changes are powerful instruments for a CAHOOTS response, with a first stage F stat of 132. A marginal CAHOOTS response to a call can be either CAHOOTS alone, or CAHOOTS as well as an officer. Columns (2) and (4) reveal that about 40% of marginal responses are CAHOOTS alone, while the remainder of them involve an officer as well. Column (3) uses an indicator for a police-only response, and reveals that 96% of compliers would receive a response from the police only if the instruments had not induced a CAHOOTS visit. Finally, column (1) shows that only 4.5% of CAHOOTS compliers would have not been responded to at all in the absence of an instrumentally-induced CAHOOTS response.

The remaining columns of Table 2 Panel A study the effects on call outcomes. A CAHOOTS response dramatically decreases the likelihood of arrest, by 20 percentage points relative to a control complier mean of 29%. Interestingly, this control complier mean is much higher than the overall sample average of 14.2%, implying that CAHOOTS is often dispatched to relatively tense situations. However, there is no statistically significant effect on the use of force (column (6)), consistent with CAHOOTS vans being dispatched only to calls with a very low likelihood of force in the first place.

CAHOOTS expansions may also free up departmental resources for other purposes. Figure 2 shows the effect of a marginal CAHOOTS response on indicators for a response occurring in less than a certain amount of time. These coefficients, which are estimated using (1) and (2), reveal that an expansion-induced CAHOOTS response had no effect on the number

of very fast responses, in particular those in less than 15 minutes. However, it increased the number of calls with a response in fewer than 40 minutes by 0.40. We view this as consistent with CAHOOTS serving as a substitute for police on lower-acuity calls, but not in emergencies. Given the importance of response time in improving call outcomes (Blanes i Vidal and Kirchmaier, 2018; DeAngelo et al., 2023), we view this as an important additional benefit of crisis response programs.

Robustness

We conduct a number of robustness checks on the difference-in-differences results. First, while our information on each call is limited, we do observe the exact address. We use this to map calls to their 2010 Census tract, allowing us to 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 rats, as well as per capita and household income. Panel B of Table 2 replicates our main difference-in-difference results while controlling for these call-level characteristics; the results are almost identical. Panel C instead uses more exacting tract-city-design fixed effects, and again finds very similar results.

Second, we estimate the reduced form analogs of (1)-(2) in Table A3. Reassuringly, each expansion increases the likelihood of a CAHOOTS response and decreases the likelihood of arrest. However, the policy changes differ substantially in how much they affect the likelihood of a CAHOOTS response, with the Eugene expansions generally having a larger effect. Given that Eugene also has a higher call volume than Springfield, this means that the IV estimates place greater weight on Eugene than Springfield.

We also estimate event study specifications. Figure A2 estimates the dynamic reduced form effects of the policy change via the following equation:

$$Y_{ijt} = \sum_{s \neq 0} \beta_s \mathbb{1}[\text{Time to change} = s] + \phi_{d(j)t} + \theta_{d(j)c(i)} + \sum_{j' \neq j} \pi_{jj'} Z_{ij'} + u_{it} \tag{3}$$

Note that because (3) does not interact the time-to-policy-change indicators with the policy change, the implicit weights on the different designs is different than in (1) and (2). We address this by reweighting the observations in (3) by the ratio of the design-average residual variance of $\widehat{D}_{ij}$ from (2) and the design-average residual variance of $\widehat{D}_{ij}$ from a first-stage regression that uses a single post-policy instrument. However, the results look nearly identical if we do not reweight.

Figure A2 shows that the likelihood of a CAHOOTS response jumps immediately following the policy changes, and although there are slightly higher CAHOOTS levels in the 2 years preceding the six months before expansion, these differences are slight in comparison to the jump in CAHOOTS likelihood falling the policy change. We see no evidence of pre-trends on the likelihood of arrest, and an immediate decline after the policy change. While most of the post-policy change coefficients are statistically insignificant by themselves, they are all very close in magnitude, motivating our static approach in (1) and (2).[5]

Finally, we assess robustness to alternative ways of constructing counterfactual outcomes. Table A5 estimates the reduced-form effect of the call coming in after an expansion using both the traditional two-way fixed effects designs analogous to (1) and (2) with 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, but the results are very similar.

4 Effects of marginal police and CAHOOTS responses

4.1 Busyness research design

The difference-in-difference estimates reveal the effect of the initial rollouts of CAHOOTS to new times of the day and geographies. Most of the added CAHOOTS dispatches would have counterfactually been responded to by the police, but after the expansion receive both a police and CAHOOTS response. In other words, in the initial rollout CAHOOTS serves largely as a complement to the police, and does not expand the universe of calls that receive a response.

In this section we study the effect of marginal expansions in CAHOOTS availability. To do so, we exploit quasi-exogeneous variation in CAHOOTS availability arising from the number and type of 911 calls that are unresolved at the time of the focal call, focusing on times and places where the initial CAHOOTS expansion has already occurred. These marginal calls are much more likely to have had no counterfactual response, suggesting that the number of calls with the scope for complementarities between the police and CAHOOTS is more limited after the initial expansion. As a result, additional CAHOOTS availability slightly lowers the likelihood of arrest and use of force, although these differences are statistically insignificant. In contrast, police dispatches resulting from unexpected availability dramatically increase arrest and use of force rates.

[5] A test of equality on the post-treatment coefficients has a p-value of 0.30 when the outcome is a CAHOOTS response, and 0.97 when the outcome is arrest.

This design is motivated by the details of how the emergency communications center in Lane County operates. At any given time, the 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 calltakers job is to talk to the caller to try to understand the situation and to then dispatch the appropriate response. 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, although they take extensive notes in the CAD system that we do not have access to. They also assign a numerical priority to the call.

Sitting in the same room as the calltakers are the dispatchers. The dispatchers job is to communicate with emergency responders and assign specific units to calls. As the calltaker talks to the caller and adds information about the call into the CAD system, the dispatcher decides what unit—the specific police, fire, ambulance, or CAHOOTS vehicle or individual—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 been responded to or it has been deemed that a response is no longer necessary, the call is closed.

Our approach is to use the number of outstanding CAHOOTS- and police-assigned calls as instruments for the response to the focal call. For each focal 911 call, we count the number of other 911 calls that have been received but not yet closed of each type (CAHOOTS or police). We use these as instruments in the following 2SLS regression that relates outcomes Y_i to the type of response:

$$Y_i = \beta_c C_i + \theta_p Z_i^p + \phi_r + u_i \tag{4}$$

$$C_i = \alpha_c Z_i^c + \gamma_p Z_i^p + \psi_r + \varepsilon_i \tag{5}$$

where C_i is a dummy for a CAHOOTS response (whether or not there is also a police response), Z_i^C and Z_i^P are the number of oustanding CAHOOTS and police calls measured in standard deviations, and ϕ_r and ψ_r 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 busyness. We also estimate an analogous specification that instruments for a police response while controlling for CAHOOTS busyness, as well as a version that instruments for both types of

response:

$$Y_i = \beta_c C_i + \beta_p P_i + \phi_r + u_i \tag{6}$$

$$C_i = \alpha_{cc} Z_i^c + \alpha_{cp} Z_i^p + \psi_r^c + \varepsilon_i \tag{7}$$

$$P_i = \alpha_{pc} Z_i^c + \alpha_{pp} Z_i^p + \psi_r^p + e_i \tag{8}$$

Note that throughout the 2SLS analysis, C_i and P_i are not mutually exclusive. The coefficients on these regressors can therefore be interpreted as the effect of adding that response, holding fixed the other response. We decompose these effects further in Section 4.4.

4.2 First stage and placebos

Figure 3 visualizes the first stages of this design. Panel A shows the effect of call busyness on the likelihood of a CAHOOTS response. The relationship is negative and nearly monotonic; an additional standard deviation of CAHOOTS call busyness (approximately 3.5 additional calls) decreases the likelihood of a CAHOOTS response by 2.1 percentage points. Panel B shows the same relationship for the police; an additional SD of police call busyness decreases the likelihood of a police response by 5.4 percentage points.

A natural concern is that call busyness is related to call characteristics even conditional on our rich set of fixed effects, perhaps through multiple calls about the same event, or through a city-level shock such as a spell of hot weather. To assuage this concern we conduct a battery of placebo tests to confirm that busyness is uncorrelated with observable characteristics.

Table 3 studies whether the busyness instruments are correlated with pre-existing characteristics of the neighborhood from which the call originates. Using the address associated with each call (which comes either from a cell phone's GPS or the address a landline is registered at) we assign the mean characteristics of the 2010 Census block group to the call. Each row of column (1) reports the coefficient on call busyness in a regression with the outcome given in the row header; for example an additional SD of call busyness predicts that the call has come from a block-group with a statistically insignificant additional 0.00155 people. Each block-group characteristic is uncorrelated with call busyness and, as revealed by the joint test at the bottom of the table, the block-group characteristics do not jointly predict busyness conditional on the city-month-day of week-hour of day fixed effects. However, these characteristics are strongly predictive of actual assignment; column (2) shows that almost all of the characteristics are related to CAHOOTS assignment and that the joint test rejects with a p-value of 0. The predictive power of the block-group characteristics for actual as-

signment suggests that our placebo exercises should have some power to detect violations of exogeneity. This is particularly clear at the bottom of the table, where the outcome is the likelihood of arrest as predicted using the block-group characteristics. A CAHOOTS response is associated with a 7.5 percentage point higher predicted arrest rate (mostly reflecting the difference between a call that is responded to at all versus not), but an additional SD of busyness increases the predicted arrest rate by a small and statistically insignificant 0.125 percentage points.

Finally, Figure 3 contains some additional placebo checks. As discussed above, this figure reveals that the busyness instruments have a large effect on the type of dispatch. As a placebo test, we use the fixed block-group characteristics in Table 3 to predict each type of treatment, and then regress predicted treatment on the instruments. The red dots and line in Figure 3 show that there is no substantive or statistically significant relationship between the instruments and predicted treatment, again consistent with the instruments being exogenous to unobserved call characteristics.

4.3 Effect of marginal dispatches on call outcomes

Panel A of Table 4 contains estimates of the effect of a busyness-induced CAHOOTS dispatch on call outcomes. We first examine the effect of a marginal CAHOOTS dispatch on the type of response. Column (1) takes an indicator for any response as the outcome; the results reveal that a slim majority—51.4%—of marginal CAHOOTS calls would not have otherwise received any response. Correspondingly, column (3) shows that 48.6% would have been police only. Of the marginal CAHOOTS calls, 30% of them also have a police response (column (4)) and the remaining 70% are CAHOOTS-only (column (2)).

Columns (5) and (6) estimate the effect of a CAHOOTS response on arrest and use of force. The likelihood of arrest declines by 5 pp relative to a control complier mean of 9.3%, although this difference is not statistically significant.

Panel B of Table 4 examines the effect of a marginal police dispatch on call outcomes. Almost all of these calls would otherwise not be responded to; column (1) shows that a marginal police dispatch increases the likelihood of any response by 89.4%. Furthermore, almost all—87%—of the marginally-dispatched calls are police-only; only 13% involve both CAHOOTS and the police. In contrast to the effect of a CAHOOTS dispatch, columns (5) and (6) show that the marginal calls are relatively likely to result in arrest or the use of force. 19.6% of marginal police calls end in an arrest and 1.5% of them with force, which is a higher rate than the average police call (see Table 1).

Finally, Panel C instruments for both a CAHOOTS response and a police response in the same regression. This regression captures a slightly different estimand, because by instrumenting for both treatments we shut down the substitution margin.[6] However, the second stage results in Table 4 are largely similar to the single-treatment models, with a CAHOOTS response slightly but insignficant reducing arrests and use of force, and a police response increasing them.

4.4 Discussion

This section has studied the effect of a CAHOOTS response that is induced by lower-than-usual busyness, and found that it moderately but insignificantly reduces arrests. This is in contrast to the expansion-induced CAHOOTS responses discussed in Section 3, which reduce arrests by nearly four times as much.

One possible reason for the differential effects of a CAHOOTS response is the differences in the counterfactual response, as well as whether the police respond along with CAHOOTS. An expansion-induced CAHOOTS response reduced the number of police-only calls by over 0.95, replacing them with CAHOOTS-only and (to a slightly greater extent) responses by both CAHOOTS and the police. In contrast, a busyness-induced CAHOOTS response decreases the number of police-only calls by only 0.326-0.486, depending on the specification. The larger effects in the expansion design are therefore most consistent with the largest gains (in terms of reductions in arrests) from CAHOOTS responses coming from them substituting the response away from being only the police.

Slightly more structure on our 2SLS regressions makes this point more clear. One way of conceptualizing our setting is that there are four possible treatments: no response, a CAHOOTS-only response, a police-only response, and a response by both types of units. We denote these responses $\{n, c, p, b\}$ respectively. In Appendix A1 we show that under the standard monotonicity condition that each instrument weakly increases the likelihood that its treatment is included—e.g., that a CAHOOTS expansion can move calls from no response to CAHOOTS-only, from police to CAHOOTS-only, and from police to both—the shares of each compliance group are identified. In particular, we identify the share of compliers where the CAHOOTS response is a complement to the police ($p \to b$), the share where the CAHOOTS response is a substitute ($p \to c$), and the share where the call would otherwise

[6] A causal interpretation under treatment effect heterogeneity requires that each instrument affects only one treatment; one implication of this is that the off-diagonal first stage coefficients in (7)-(8) should be equal to zero. Table A6 shows that the off-diagonal coefficients are small and a formal test of the null hypothesis has a p-value of 0.051, suggesting that any violations are slight.

not have been responded to ($n \to c$).

Table A7 reports these shares, and reveals a sharp contrast between the compliance groups in the initial expansions versus the marginal respondents to the busyness instruments. Each marginal CAHOOTS response in the initial expansion added 0.6 calls where CAHOOTS is a complement to the police, and 0.36 where they are a substitute. Only 0.045 calls would not otherwise be responded to. In contrast, for the busyness-induced CAHOOTS responses, 51.4% of calls would not have been responded to, with sharp declines in the number of complement and substitute compliers.

We view these results as suggesting that only a limited degree of substitution between CAHOOTS and the police is possible. The initial expansions boost the share of responses that include CAHOOTS by 8 percentage points from a base of nearly zero (see Figure A2), but as discussed above, the marginal calls after the expansions are much more likely to be counterfactually not responded to. This implies that even at current levels of CAHOOTS response (11% of 911 calls in 2019, the last pre-pandemic year), there has been a decline in the extent to which CAHOOTS can act as a substitute for the police. Instead, additional CAHOOTS expansions are likely to provide services to callers who would otherwise receive no response.

4.5 Decomposing the effect of CAHOOTS

As discussed above, our IV estimates can be interpereted as a weighted average of three compliance-group specific effects: the effect of CAHOOTS relative to the police ($p \to c$), of both CAHOOTS and the police relative to the police only ($p \to b$), and CAHOOTS relative to no response ($n \to c$). These effects can be separately identified, however, with some additional assumptions on treatment effect heterogeneity. Specifically, as we show in Section A1.3, if we assume that the treatment effects (but not the size of each compliance group) are constant across the five different difference-in-differences designs, this is sufficient to identify compliance-group-specific treatment effects. Intuitively, this works because the IV estimate in each design is equal to the share of each compliance group (which is identified from the data) multiplied by its compliance-group-specific treatment effect. Since there are five designs but only three compliance groups, the set of equations is overidentified and has a unique solution.

Table A8 contains the results for the effects on arrest likelihood. Column (1) shows that the $p \to c$ effect is about twice the size of the $p \to b$ effect (-0.171 and -0.286, respectively), indicating that substituting CAHOOTS for the police has a larger effect on arrests than

complementing the police with CAHOOTS. We also find that a CAHOOTS dispatch rather than no response has no statistically significant effect on arrest rates, which we view as a validity check on our model. More formally, a Sargan-Hansen J test of the overidentifying restrictions fails to reject the null.

However, the estimates are relatively imprecise. In column (2), we impose that the $n \to c$ effect on arrests is zero, which reduces the SEs for the other coefficients by about 40%. However, the point estimates are barely affected, resulting in a p-value for the $n \to c$ effect of 0.057 and a smaller and statistically insignificant $p \to c$ effect.

In columns (3) and (4) we additionally use the busyness IV estimates from Section 4.3 as an additional moment to match. Reassuringly, the estimates are nearly unchanged and we continue to pass the overidentification tests, suggesting that our model of homogeneous compliance-group-specific treatment effects across designs is reasonable. We conclude that the largest reductions in arrests come from substituting the police with CAHOOTS, but that there may also be gains to complementing the police with CAHOOTS.

5 The long-term effect of CAHOOTS

Our analysis up until now has revealed that dispatching CAHOOTS to a call rather than the police 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 arrested those individuals likely to require a future emergency response.

In this section we study these issues by examining the long-term effects of CAHOOTS and police dispatches on future calls from the same address. While we lack individual-level information on the caller such as phone number or name, we do know the address each call is received from. We interpret future 911 calls from the same address as indicative of an unresolved problem, and thus that an effective emergency response will reduce the future calls. For each day relative to the focal call, we measure whether or not there has been a subsequent 911 call from the same address, excluding those calls within the first 24 hours to avoid capturing anything related to the focal call itself.

We begin by studying the effect of a CAHOOTS response. Panel A of Figure 4 uses (1)-(2) and (4)-(5) to estimate the effect of a CAHOOTS response on the next call from that address being received in fewer than t days for each t from 2 to 90. We find that a CAHOOTS response induced by the initial expansion reduces the likelihood of a future call by 17.6

percentage points in the first week. In the longer term, as a higher share of observations have received a subsequent call, the effect fades, but even as late as 30 days after the focal call, a CAHOOTS response still reduces the likelihood of a callback by 15 percentage points ($p = 0.16$).

We also find that a CAHOOTS response induced by capacity lowers the likelihood of a future call. Consistent with the discussion in Section 4.4 of the different margins that these responses operate on, the effect is somewhat smaller and shorter-lasting: a CAHOOTS response reduces the likelihood of a future call by 11 percentage points after a week ($p = 0.077$) but has no statistically significant effect after two weeks.

Panel B of Figure 4 studies the effect of a police response using the busyness instrument. In contrast to the effect of a CAHOOTS response, a police response slightly *increases* the likelihood of a subsequent call over the two weeks following the focal call, although these effects are significant only at the 10% level.[7] It thus appears that concerns about CAHOOTS resulting in more long-term problems are unfounded, and if anything a CAHOOTS response is an effective medium-term solution.

Robustness

An alternative measure of future calls is the number of calls over the next t days. Figure A3 shows that using this measure leads to similar conclusions as using an indicator for any call: a CAHOOTS response induced by additional capacity reduces the number of future calls, with the largest and most statistically significant reductions occurring two weeks after the focal event. A police response, however, modestly but insignificantly increases the number of calls over the first several weeks.

Note that we do not report estimates of the effect of an expansion-induced CAHOOTS response on the number of future calls. This is because the expansion instrument violates exclusion when the outcome is number of future calls: when the focal call occurs after the expansion, so do any subsequent calls, which will further decrease the likelihood of a future 911 calls. In contrast, when we are using the busyness instrument, there is no such correlation between the focal and future type of response.

[7] We reject equality of the CAHOOTS and police effect on callback within a week with a p-value of 0.037.

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.[8] 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 the labor cost of two new police officers ($72.33 per hour).[9]

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 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 10% of calls, most of the calls where CAHOOTS can

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

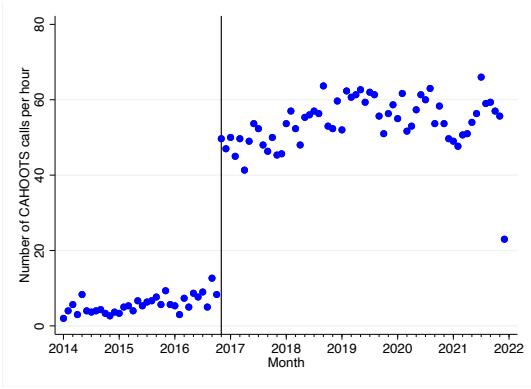
[9] In FY 2022, the city of Springfield paid $459,500 for CAHOOTS services, with 6,353 responses in 2021. This translates to a cost of $72.33 per response. Source: https://springfield-or.gov/wp-content/uploads/2022/06/FY22-CITY-Adopted-Budget_6.15.22.pdf

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 are more general limits to the crisis responder model.

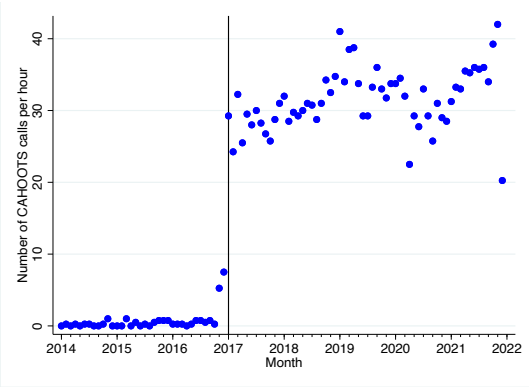
Figures

Figure 1: CAHOOTS expansions in Eugene and Springfield

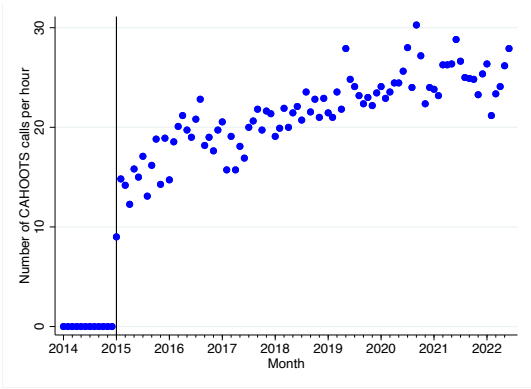
(a) Eugene 7-10 a.m.



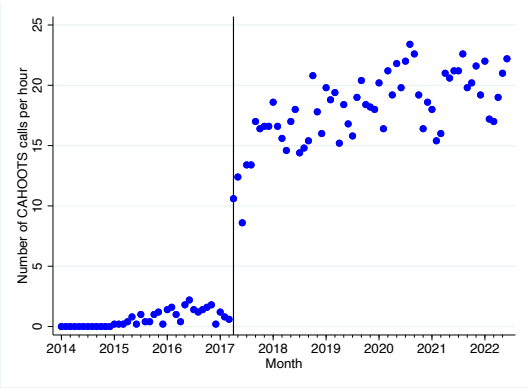
(b) Eugene 3-7 a.m.



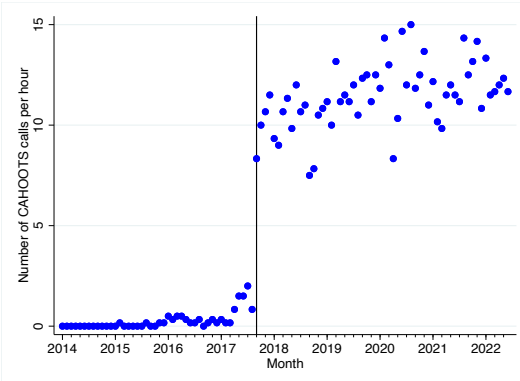
(c) Springfield 12 p.m. to 11 p.m.



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

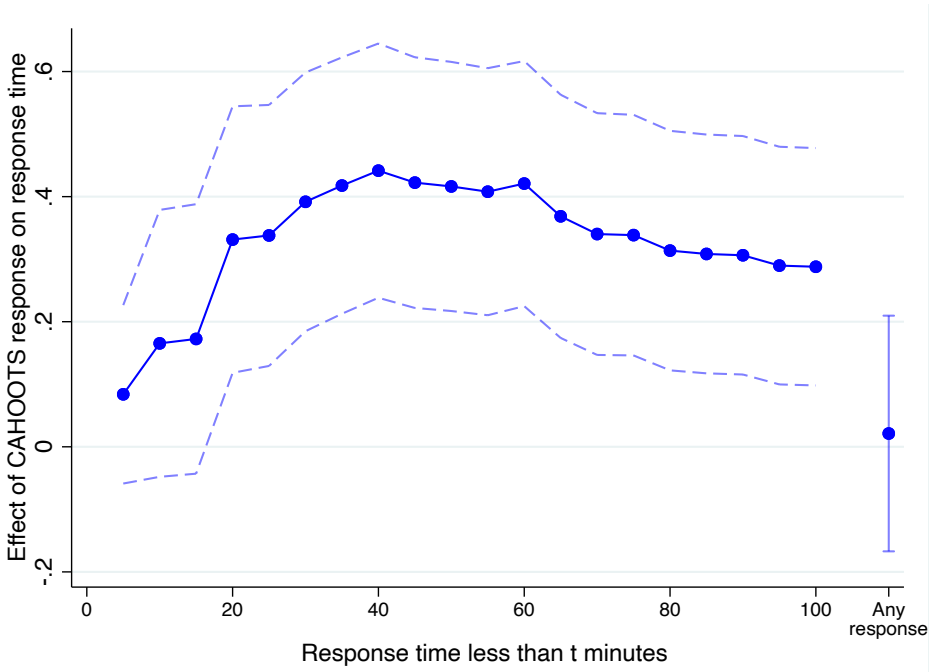


(e) Springfield 3-10 a.m.



Notes: This figure shows the changes in the number of calls that CAHOOTS is dispatched to 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 to, normalized to be measured in terms of calls per hour-month.

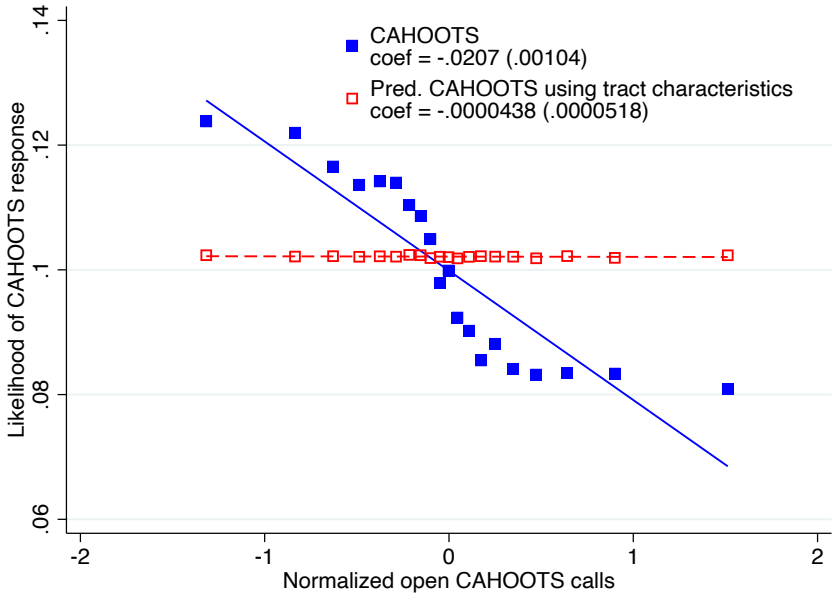
Figure 2: The impact of an expansion-induced CAHOOTS response on call response times



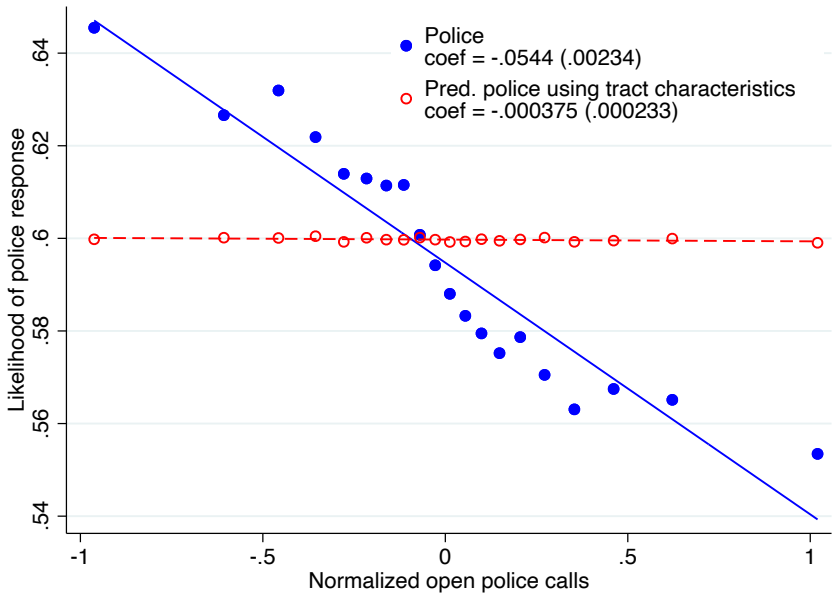
Notes: This figure shows the effect of a CAHOOTS response on the likelihood a 911 calls gets a response within a given period of time, and at all. Each point is a coefficient from a 2SLS regression of an indicator for a 911 call response (either CAHOOTS or police) occurring in fewer than t minutes from the focal call on CAHOOTS response, where CAHOOTS response is instrumented using indicators for each CAHOOTS expansion as detailed in (1)-(2) The specification includes city-design and month-design FEs, as well as controls for other policy changes. The rightmost coefficient is from a regression where the outcome variable is an indicator for any 911 call response occurring at any time. 95% confidence intervals clustered at the call level are shown in the dotted lines and whiskers.

Figure 3: First stage of busyness instruments on type of dispatch

(a) CAHOOTS response



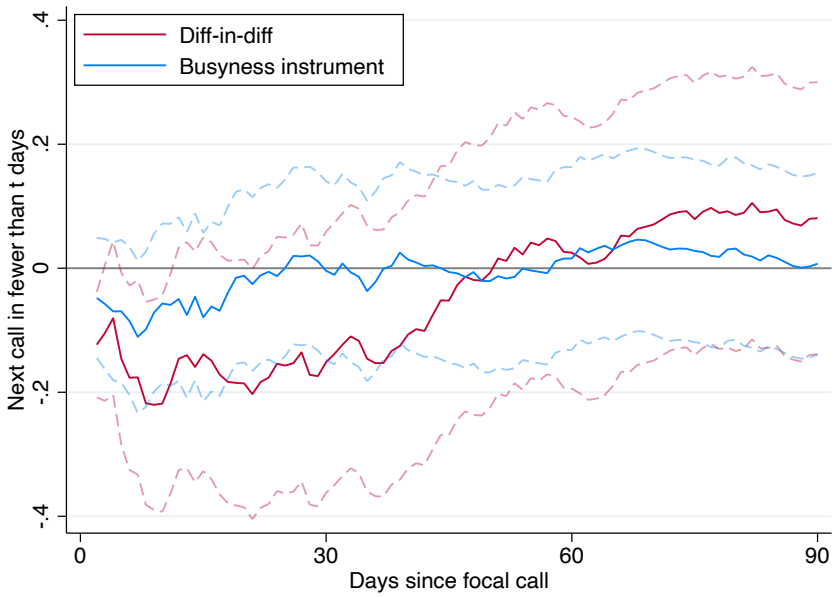
(b) Police response



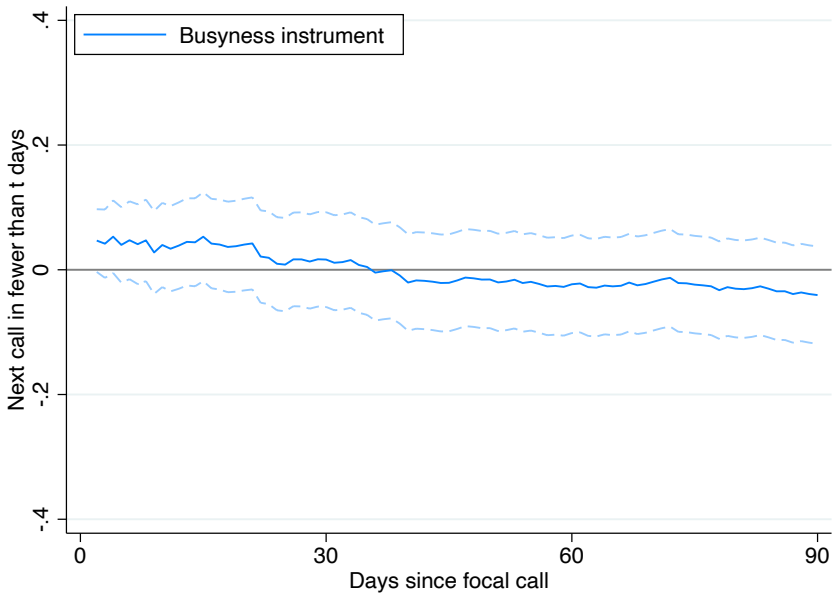
Notes: This figure shows the first-stage relationship between our standardized busyness instruments for CAHOOTS (Panel A) and police (Panel B) and the likelihood of getting that type of response (blue line) and the predicted probability of getting that response based on census tract characteristics (red line).

Figure 4: Effect of emergency response on future call from same address in fewer than t days

(a) Effect of CAHOOTS response



(b) Effect of police response



Notes: This figure shows the effect of CAHOOTS or police response on future 911 calls from the same address within a period of time up to 90 days. Panel A shows the effects of CAHOOTS response from both designs. 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 busyness instrument. Panel B shows the coefficients from a 2SLS regression of an indicator for a future 911 call within t days on police response, where police response is instrumented for using the busyness instrument. 95% confidence intervals clustered at the individual level are shown in dotted lines.

Tables

Table 1: Call characteristics and outcomes by response

	(1) No response	(2) CAHOOTS only	(3) Police only	(4) Both
Panel A: Call type				
Welfare check	0.040	0.413	0.037	0.133
Suicidal person	0.007	0.089	0.011	0.152
Traffic/driver issue	0.179	0.053	0.089	0.035
Disorderly subject	0.048	0.064	0.093	0.158
Intoxicated person	0.003	0.042	0.003	0.011
Suspicous person	0.045	0.016	0.079	0.046
Subject down	0.005	0.014	0.002	0.007
Overdose	0.016	0.006	0.010	0.027
Mischief/trespassing	0.050	0.006	0.057	0.034
Dispute	0.049	0.004	0.258	0.157
Assault	0.007	0.001	0.031	0.028
Harassment	0.011	0.001	0.020	0.007
Burglary/theft	0.048	0.001	0.045	0.010
Drunk driver	0.067	0.000	0.024	0.003
Shots fired	0.013	0.000	0.015	0.001
Armed subject	0.002	0.000	0.009	0.004
Fight	0.002	0.000	0.010	0.002
Transportation	0.003	0.041	0.000	0.006
Assist police department	0.013	0.177	0.002	0.033
Mental issue	0.002	0.017	0.009	0.015
Assist fire department	0.006	0.044	0.006	0.011
Cardiac arrest	0.012	0.001	0.006	0.018
Robbery	0.001	0.000	0.005	0.002
Warrant	0.000	0.000	0.008	0.005
Incomplete call	0.105	0.001	0.021	0.006
Other	0.267	0.022	0.151	0.094
Panel B: Predicted call responses and outcomes				
Predicted CAHOOTS response	0.066	0.414	0.067	0.217
Predicted police response	0.518	0.444	0.681	0.638
Predicted arrest	0.069	0.058	0.103	0.094
Predicted use of force	0.003	0.003	0.006	0.005
Panel C: Call outcomes				
Time until dispatch (minutes)	5.732	30.967	13.725	10.815
Time until arrival (minutes)	0.208	39.044	19.030	17.412
Arrest	0.001	0.000	0.142	0.141
Use of force	0.000	0.000	0.007	0.011
Observations	81,095	13,079	128,081	10,127
Total	232,382			

Table 2: Effects of CAHOOTS response on call outcomes using expansions as instruments

	Call responses				Call outcomes	
	Any response	Only CAHOOTS	Only police	Both	Arrest	Use of force
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Baseline specification						
CAHOOTS response	0.045	0.404***	-0.955***	0.596***	-0.199***	0.002
	(0.098)	(0.023)	(0.098)	(0.023)	(0.074)	(0.020)
Control complier mean	0.955	0	0.955	0	0.294	0.012
First-stage F-stat	131.92	131.92	131.92	131.92	131.92	131.92
Panel B: with tract controls						
CAHOOTS response	0.017	0.398***	-0.983***	0.602***	-0.218***	0.003
	(0.101)	(0.024)	(0.101)	(0.024)	(0.076)	(0.021)
Control complier mean	0.983	0	0.983	0	0.311	0.011
First-stage F-stat	121.88	121.88	121.88	121.88	121.88	121.88
Equality with Panel A (p)	0.060	0.232	0.060	0.232	0.034	0.648
Panel C: with tract FEs						
CAHOOTS response	0.087	0.406***	-0.913***	0.594***	-0.186**	0.003
	(0.097)	(0.023)	(0.097)	(0.023)	(0.073)	(0.020)
Control complier mean	0.955	0	0.955	0	0.294	0.012
First-stage F-stat	131.92	131.92	131.92	131.92	131.92	131.92
Equality with Panel A (p)	0.003	0.258	0.003	0.258	0.015	0.231
Observations	284,487	284,487	284,487	284,487	284,487	284,487

Notes: This table shows the estimated impact of a CAHOOTS response on response types and call outcomes. Each column reports the results from a 2SLS regression of the outcome on CAHOOTS response, where CAHOOTS response is instrumented using indicators for each CAHOOTS expansion. The baseline specification in Panel A includes city-design and month-design FEs and controls for other policy changes. Panel B adds controls for characteristics of the call's census tract. Panel C controls for census tract fixed effects. Standard errors clustered at the call level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Placebo tests for busyness instruments

	(1) CAHOOTS busyness	(2) CAHOOTS response	(3) Police busyness	(4) Police response
Total population	0.00155	-0.0196***	-0.00121	0.00345
	(0.00360)	(0.00757)	(0.00495)	(0.00459)
Male	0.000363	0.0785***	-0.00817*	0.0529***
	(0.00330)	(0.00725)	(0.00481)	(0.00451)
Hispanic	0.000663	0.00112	-0.0103**	0.0930***
	(0.00353)	(0.00723)	(0.00490)	(0.00456)
Black (2018)	0.000744	0.0476***	-0.000789	0.0340***
	(0.00328)	(0.00670)	(0.00460)	(0.00429)
Employed (2018)	0.00476	-0.0333***	0.00338	-0.00484
	(0.00352)	(0.00747)	(0.00495)	(0.00464)
Unemployed	-0.00352	0.0437***	-0.00541	0.0193***
	(0.00329)	(0.00708)	(0.00469)	(0.00437)
Household income (2018)	0.000245	-0.100***	-0.00622	-0.0413***
	(0.00363)	(0.00759)	(0.00505)	(0.00465)
Income (per capita)	-0.00104	-0.0773***	0.00486	-0.0722***
	(0.00365)	(0.00728)	(0.00500)	(0.00460)
Aggregate income	0.00127	-0.0801***	0.00345	-0.0554***
	(0.00365)	(0.00726)	(0.00502)	(0.00458)
Median gross rent	0.00437	-0.120***	-0.00628	-0.0634***
	(0.00352)	(0.00716)	(0.00484)	(0.00441)
Median gross rent/income	-0.000465	0.0258***	-0.000356	-0.00346
	(0.00345)	(0.00703)	(0.00482)	(0.00440)
Median home value	-0.00194	-0.00489	0.00697	-0.0945***
	(0.00337)	(0.00676)	(0.00462)	(0.00423)
Predicted arrests	0.00125	0.0747***	-0.00509	0.0868***
	(0.00312)	(0.00626)	(0.00428)	(0.00392)
Predicted use of force	0.000676	0.00410	-0.00290	0.0398***
	(0.00288)	(0.00599)	(0.00402)	(0.00374)
Joint F	0.737	21.517	1.588	56.003
p-value	0.728	0.000	0.080	0.000
Observations	232,382	232,382	232,382	232,382

Notes: This table tests whether the busyness instruments are correlated with pre-existing characteristics of the neighborhood from which the call originates. Each cell reports the coefficient from a regression of the outcome in the row header on the covariate in the column header and call month by day of week by hour of day by city fixed effects. Moving from left to right, the columns show the relationships between our CAHOOTS busyness instrument, CAHOOTS assignment to a call, our police busyness instrument, and police assignment to a call. Robust standard errors in parentheses. Asterisks based on conventional inference: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Effects of CAHOOTS and police dispatch on call outcomes

	Call response				Call outcomes	
	Any	Only CAHOOTS	Only police	Both	Arrest	Force
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Instrumenting for CAHOOTS						
CAHOOTS response	0.514***	0.697***	-0.486***	0.303***	-0.052	-0.005
	(0.077)	(0.027)	(0.077)	(0.027)	(0.044)	(0.009)
Control complier mean	.486	0	.486	0	.0927	.00618
Observations	232,382	232,382	232,382	232,382	232,382	232,382
Panel B: Instrumenting for police						
Police response	0.894***	-0.106***	0.870***	0.130***	0.196***	0.015***
	(0.020)	(0.020)	(0.017)	(0.017)	(0.023)	(0.005)
Control complier mean	.106	.106	0	0	.00244	-.0000579
Observations	232,382	232,382	232,382	232,382	232,382	232,382
Panel C: Instrumenting for both responses						
CAHOOTS response	0.674***	0.674***	-0.326***	0.326***	-0.016	-0.003
	(0.025)	(0.025)	(0.025)	(0.025)	(0.043)	(0.009)
Police response	0.877***	-0.123***	0.877***	0.123***	0.196***	0.015***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.023)	(0.005)
Observations	232,382	232,382	232,382	232,382	232,382	232,382

Notes: This table reports coefficients from a 2SLS regression of the outcome on CAHOOTS or police response. In Panels A and B, CAHOOTS (police) response is instrumented for using the number of open CAHOOTS (police) calls, measured in standard deviations and controlling for the other instrument. In Panel C, both types of response are jointly instrumented for using the number of open CAHOOTS calls and open police calls. Call month by day of week by hour of day by city fixed effects are included in all regressions. Standard errors clustered at the individual level are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

References

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.

Bhuller, M. and H. Sigstad (2024). 2sls with multiple treatments. *Journal of Econometrics 242*(1), 105785.

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.

DeAngelo, G., M. Toger, and S. Weisburd (2023). Police response time and injury outcomes. *The Economic Journal 133*, 2147–2177.

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

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

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

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

Frandsen, B., L. Lefgren, and E. Leslie (2023). Judging judge fixed effects. *American Economic Review 113*(1), 253–277.

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.

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.

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.

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

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.

A1 Econometric Appendix

In this section, we decompose the 2SLS estimands into their constituent subLATES. The shares of each compliance group—for example, compliers who are moved from no response to a CAHOOTS-only response—are identified, but the effects for each group are typically not.

We begin by splitting the responses to a 911 call into four groups: no response (n), only CAHOOTS (c), only the police (p), and both CAHOOTS and the police (b). Let $D_i \in \{n, c, p, b\}$ denote the response to call i. Let $Y_i(d)$ denote the potential outcome of interest under treatment regime d. To tractably describe the different counterfacuals we will focus on the case of binary instruments, $Z_i^c \in \{0, 1\}$ that shifts the likelihood of a CAHOOTS response and $Z_i^p \in \{0, 1\}$ that shifts the likelihoof of a police response. For a given instrument, there are 16 potential *response types*, in the terminology of Heckman and Pinto (2018).

Guided by our institutional setting, we make two key assumptions. First, we assume that the CAHOOTS (police) instrument increases the likelihood of a response that involves CAHOOTS (the police) for each individual, following the monotonicity assumption or Imbens and Angrist (1994).[1] Second, we assume that conditional on the other instrument, each instrument does not affect the likelihood of receiving the other treatment, except in the case of substitution. For example, Z_i^c can move calls from p to c, but not from p to n.

Table A1 illustrates how these restriction reduce the set of possible response types from 16 to 7. The blue color cells refer to response types that have been eliminated due to the standard monotonicity assumption and the red color cells to transitions that have been eliminated due to the no cross effects assumption.[2]

Table A1: Response types

		Z_i^c=1			
		n	c	p	b
Z_i^c=0	n	π_{nn}	π_{nc}	-	-
	c	-	π_{cc}	-	-
	p	-	π_{pc}	π_{pp}	π_{pb}
	b	-	-	-	π_{bb}

A. Instrument that increases CAHOOTS (Z_i^c)

		Z_i^p=1			
		n	c	p	b
Z_i^p=0	n	π_{nn}	-	π_{np}	-
	c	-	π_{cc}	π_{cp}	π_{cb}
	p	-	-	π_{pp}	-
	b	-	-	-	π_{bb}

B. Instrument that increases police (Z_i^p)

Notes: This figure reports the response types allowed under each instrument, holding the value of the other instrument fixed. $\pi_{xy} \equiv Pr[D(0)=x, D(1)=y]$ for the relevant instrument.

Next we define the two monotonicity restriction more formally. Our monotonicity restric-

[1] Note that we can relax the Imbens and Angrist (1994) monotonicity to an average monotonicity style assumption following the work of Frandsen et al. (2023) and its generalization to the case of multiple treatments by Bhuller and Sigstad (2024).

[2] This assumption is similar but distinct from the no cross effects assumption of Bhuller and Sigstad (2024), who focus on a setting with three mutually exclusive treatments.

tions, which are all conditional on the other instrument, can be stated as:

$$\mathbb{1}[D_i(Z_i^c{=}1, Z_i^p) \in \{c, b\}] \geq \mathbb{1}[D_i(Z_i^c{=}0, Z_i^p) \in \{c, b\}] \tag{A1}$$
$$\mathbb{1}[D_i(Z_i^c, Z_i^p{=}1) \in \{p, b\}] \geq \mathbb{1}[D_i(Z_i^c, Z_i^p{=}0) \in \{p, b\}] \tag{A2}$$

Our no cross effects assumption can be stated as:

$$\mathbb{1}[D_i(Z_i^c{=}1, Z_i^p) \in \{c, p, b\}] \geq \mathbb{1}[D_i(Z_i^c{=}0, Z_i^p) \in \{c, p, b\}] \tag{A3}$$
$$\mathbb{1}[D_i(Z_i^c{=}1, Z_i^p) \in \{p, b\}] \leq \mathbb{1}[D_i(Z_i^c{=}0, Z_i^p) \in \{p, b\}] \tag{A4}$$
$$\mathbb{1}[D_i(Z_i^c{=}1, Z_i^p) \in \{b\}] \geq \mathbb{1}[D_i(Z_i^c{=}0, Z_i^p) \in \{b\}] \tag{A5}$$
$$\mathbb{1}[D_i(Z_i^c, Z_i^p{=}1) \in \{c, p, b\}] \geq \mathbb{1}[D_i(Z_i^c, Z_i^p{=}0) \in \{c, p, b\}] \tag{A6}$$
$$\mathbb{1}[D_i(Z_i^c, Z_i^p{=}1) \in \{c, b\}] \leq \mathbb{1}[D_i(Z_i^c, Z_i^p{=}0) \in \{c, b\}] \tag{A7}$$
$$\mathbb{1}[D_i(Z_i^c, Z_i^p{=}1) \in \{b\}] \geq \mathbb{1}[D_i(Z_i^c, Z_i^p{=}0) \in \{b\}] \tag{A8}$$

A1.1 What estimand do Wald estimators recover?

These assumptions allow us to characterize what we can learn from IV regressions in this environment. We focus on a Wald estimator of the effect of any type of CAHOOTS response, $\mathbb{1}[D \in \{c, b\}]$, and show that the estimand is composed of three factors: (1) *complementarity* of CAHOOTS to police; (2) *substitution* from police to CAHOOTS; and (3) the impact of *increasing services* to calls the police would respond otherwise:

$$\frac{E[Y|Z^c{=}1] - E[Y|Z^c{=}0]}{E[\underbrace{1(D \in \{c, b\})}_{\text{CAHOOTS response}}|Z^c{=}1] - E[1(D \in \{c, b\})|Z^c{=}0]} \tag{A9}$$

$$= \underbrace{\frac{\pi_{pb}}{\pi_{nc} + \pi_{pc} + \pi_{pb}}}_{\equiv\, s_{pb}} \underbrace{E[Y(b) - Y(p)|D_i(1){=}b, D_i(0){=}p]}_{\substack{\text{Adding CAHOOTS to a call} \\ \text{only police would have responded}}} \tag{A10}$$

$$+ \underbrace{\frac{\pi_{pc}}{\pi_{nc} + \pi_{pc} + \pi_{pb}}}_{\equiv\, s_{pc}} \underbrace{E[Y(c) - Y(p)|D_i(1){=}c, D_i(0){=}p]}_{\text{Substituting from police to CAHOOTS}}$$

$$+ \underbrace{\frac{\pi_{nc}}{\pi_{nc} + \pi_{pc} + \pi_{pb}}}_{\equiv\, s_{nc}} \underbrace{E[Y(c) - Y(n)|D_i(1){=}c, D_i(0){=}n]}_{\text{CAHOOTS response to a call the police would not}}$$

The estimand in (A9) recovers a well-defined causal parameter when all three effects are present. However, for some outcomes, like arrests and use of force, this expression will be even simpler. This is because these outcomes are equal to zero under both n and c, and so the last expression can be eliminated.

The interpretation of the Wald estimator using an instrument that only shifts police response

is simpler. It is well-defined for all outcomes and it recovers:

$$\frac{E[Y|Z^p{=}1] - E[Y|Z^p{=}0]}{E[\underbrace{1(D \in \{p,b\})}_{\text{Any police}}|Z^c{=}1] - E[1(D \in \{p,b\})|Z^c{=}0]} \tag{A11}$$

$$= \frac{\pi_{cb}}{\pi_{np}+\pi_{cp}+\pi_{cb}} \underbrace{E[Y(b) - Y(c)|D_i(1){=}b, D_i(0){=}c]}_{\substack{\text{Adding police to a call} \\ \text{only CAHOOTS would have responded}}} \tag{A12}$$

$$+ \frac{\pi_{cp}}{\pi_{np}+\pi_{cp}+\pi_{cb}} \underbrace{E[Y(p) - Y(c)|D_i(1){=}p, D_i(0){=}c]}_{\text{Substituting from CAHOOTS to police}}$$

$$+ \frac{\pi_{np}}{\pi_{np}+\pi_{cp}+\pi_{cb}} \underbrace{E[Y(p) - Y(n)|D_i(1){=}p, D_i(0){=}n]}_{\substack{\text{Police response to a call} \\ \text{that would not get service otherwise}}}$$

A1.2 Identifying the size of each compliance group

While the treatment effects for each compliance group are not identified without further assumptions, the size of each group is. This can be seen in the following derivation:

$$E[\mathbb{1}[D_i{=}b]|Z_i^c{=}1] - E[\mathbb{1}[D_i{=}b]|Z_i^c{=}0] = \pi_{pb} \tag{A13}$$

$$E[\mathbb{1}[D_i{\in}\{n,c\}]|Z_i^c{=}1] - E[\mathbb{1}[D_i{\in}\{n,c\}]|Z_i^c{=}0] = \pi_{pc} \tag{A14}$$

$$E[\mathbb{1}[D_i{\in}\{c,p,b\}]|Z_i^c{=}1] - E[\mathbb{1}[D_i{\in}\{c,p,b\}]|Z_i^c{=}0] = \pi_{nc} \tag{A15}$$

These estimates can be rescaled by the first-stage effect of the instrument on any CAHOOTS response to recover the share of compliers in each of the three compliance groups.

Similarly for the police instrument, we have

$$E[\mathbb{1}[D_i{=}b]|Z_i^p{=}1] - E[\mathbb{1}[D_i{=}b]|Z_i^p{=}0] = \pi_{cb} \tag{A16}$$

$$E[\mathbb{1}[D_i{\in}\{n,p\}]|Z_i^p{=}1] - E[\mathbb{1}[D_i{\in}\{n,p\}]|Z_i^p{=}0] = \pi_{cp} \tag{A17}$$

$$E[\mathbb{1}[D_i{\in}\{c,p,b\}]|Z_i^p{=}1] - E[\mathbb{1}[D_i{\in}\{c,p,b\}]|Z_i^p{=}0] = \pi_{np} \tag{A18}$$

While these derivations use the binary-instrument Wald estimand, in the paper we use the full support of the instrument to estimate the shares of each compliance group, ensuring that the shares correspond to our reported estimates. For example, to estimate the share of compliers who move from police-only to both police and CAHOOTS as a result of an expansion-induced CAHOOTS response, we estimate (1)-(2) with the outcome $\mathbb{1}[D_i{=}b]$.

A1.3 Estimating subLATEs by combining design-specific estimates

We can also use the different designs to estimate the relative effect of each treatment. In particular, we assume that the effect is homogenous across designs within the three com-

pliance groups present under our IV assumptions: $n \to c$ compliers, $p \to c$ compliers, and $p \to b$ compliers. Then, in each design d the treatment effect β_d can be decomposed as

$$\beta_d = s^d_{pb}\theta_{pb} + s^d_{pc}\theta_{pc} + s^d_{nc}\theta_{nc} \tag{A19}$$

where s are the compliance group shares in (A10). Stacking this expression into matrix notation, it can be more parsimoniously expressed as $\beta = S\theta$, where β is $(d \times 1)$, S is $(d \times 3)$ and θ is (3×1). We use minimum distance to estimate

$$\widehat{\theta} = \underset{\theta}{\text{argmin}}\ (\widehat{\beta} - S\theta)'W(\widehat{\beta} - S\theta) \tag{A20}$$

We pick $W = diag(V[\widehat{\beta}])^{-1}$, implicitly restricting the estimates to be uncorrelated across designs. Then, we have that

$$V(\widehat{\theta}) = (S'WS)^{-1}(S'WV(\widehat{\beta})WS)(S'WS)^{-1} \tag{A21}$$

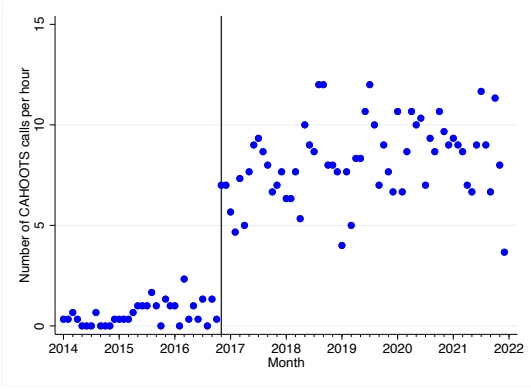
$$= (S'WS)^{-1} \tag{A22}$$

Also note that the objective $Q(\widehat{\theta})$ is distributed $\chi^2(d-q)$ under the null, where d is the number of designs and q is the number of parameters. This allows us to construct a test of the overidentifying restrictions.

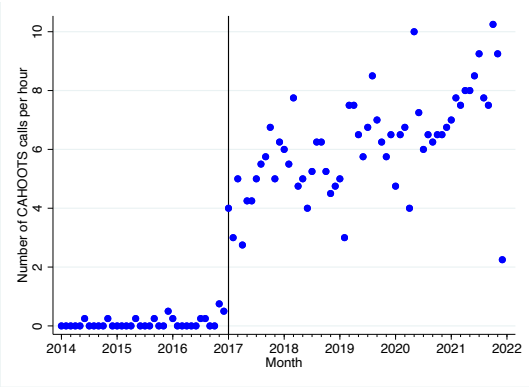
A2 Appendix Figures

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

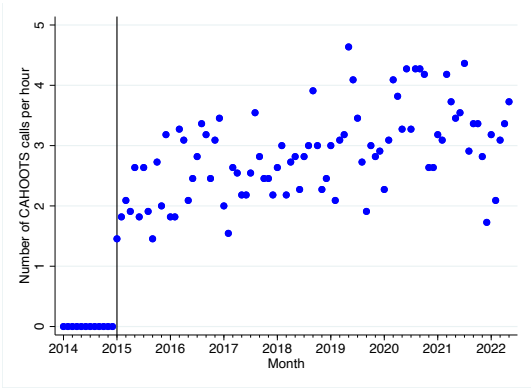
(a) Eugene 7-10 a.m.



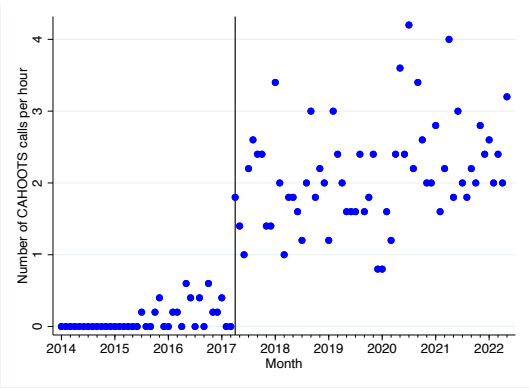
(b) Eugene 3-7 a.m.



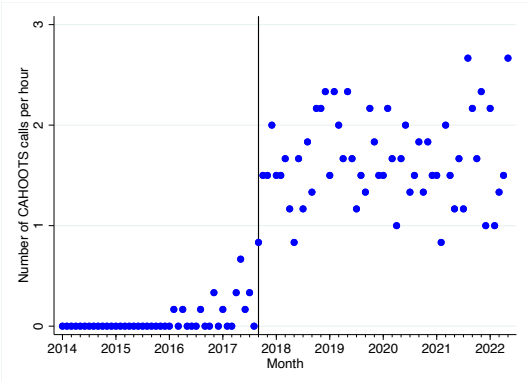
(c) Springfield 12 p.m. to 11 p.m.



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



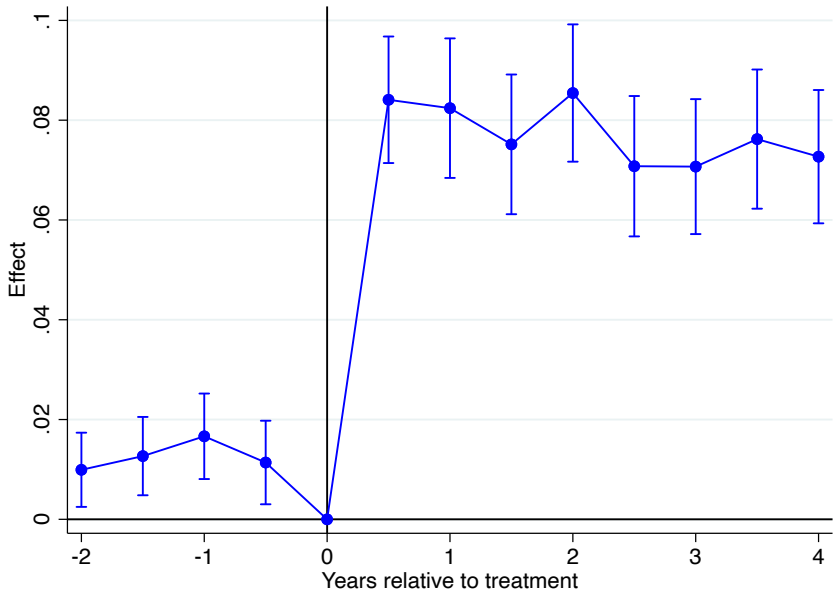
(e) Springfield 3-10 a.m.



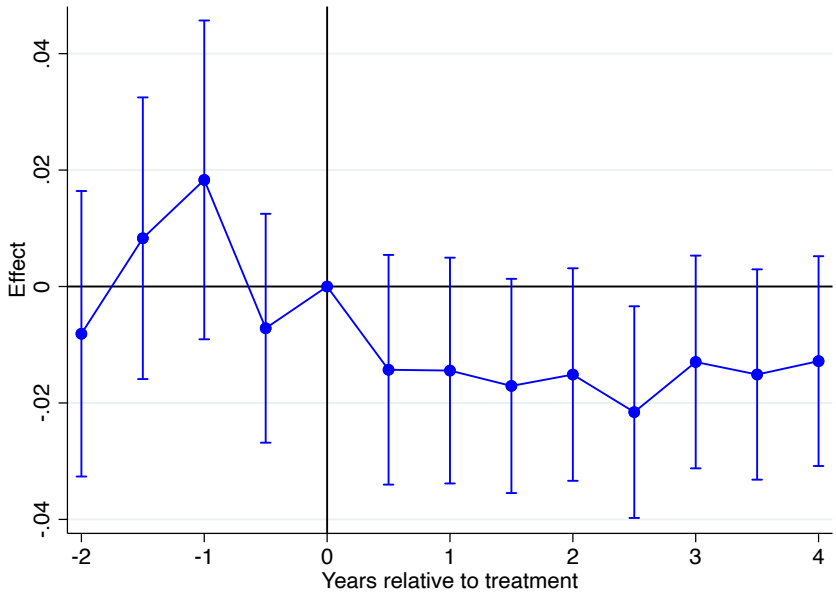
Notes: This figure shows the changes in the number of calls that CAHOOTS is dispatched to as services are expanded (Eugene) or introduced (Springfield). Panels A and B report expansions in Eugene and Panels C, D, and E expansions in Springfield. Each point in the figures reports the number of 911 calls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

Figure A2: Event study specifications

(a) CAHOOTS response

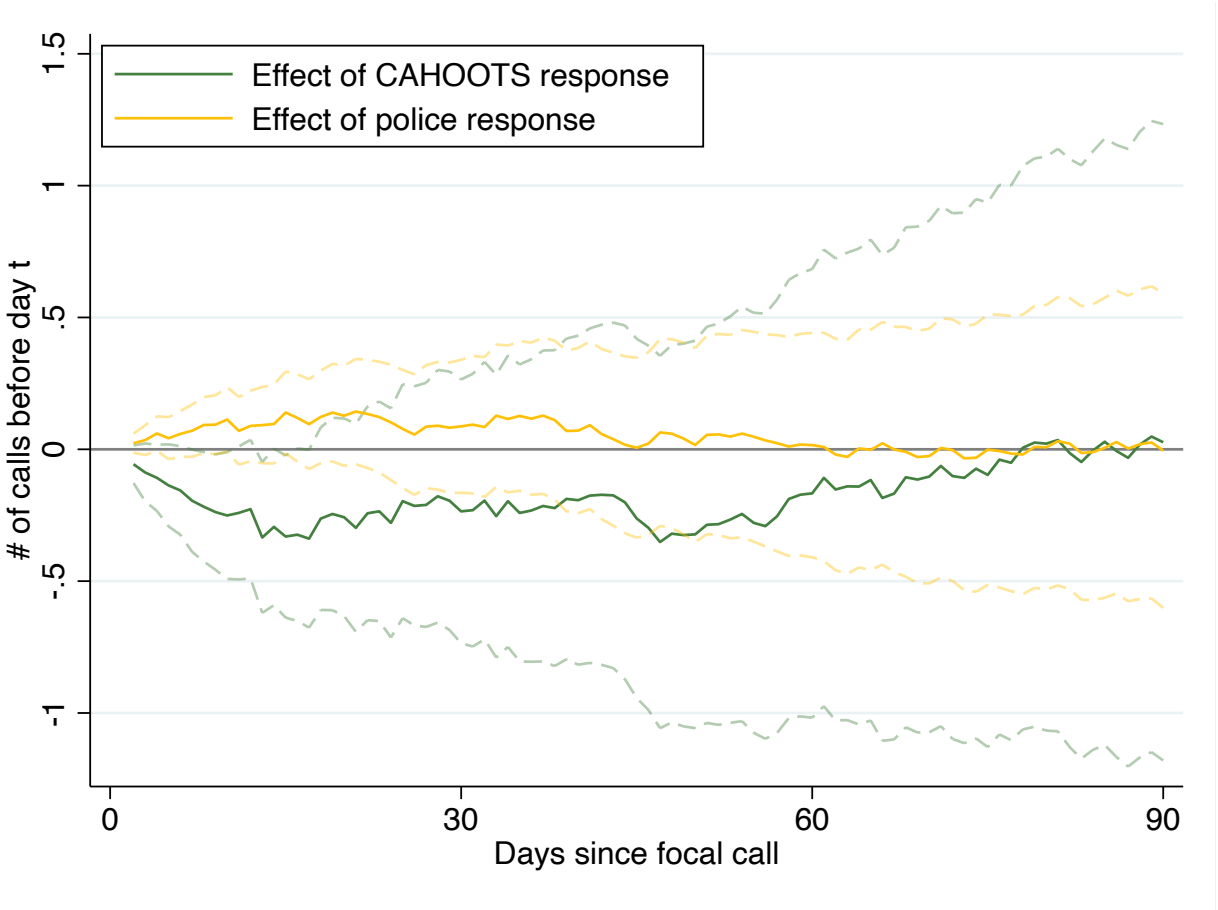


(b) Arrest



Notes: This figure shows the effect of CAHOOTS expansions on CAHOOTS responses (panel A) and arrests (Panel B) in six month windows. Estimates are from a regression of the outcome in each panel on indicators for CAHOOTS expansions interacted with indicators for each of the six month windows and fixed effects for design by month, design by city, and for other policy changes. We plot the coefficients on the interactions between six month group and CAHOOTS expansions. The coefficient on the interaction in the first six months before an expansion normalized to zero. 95% confidence intervals clustered at the call level shown in whiskers.

Figure A3: Effect of emergency response on number of calls before day t



Notes: This figure shows the effect of CAHOOTS or police response an indicator for a future 911 call from the same address within a period of time up to 90 days. Each coefficient comes from a 2SLS regression of an indicator for a future call within t days on CAHOOTS/police response. CAHOOTS/police response is instrumented using the buysness design. 95% confidence intervals clustered at the call level in dotted lines.

A3 Appendix Tables

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 dates that each of the 5 CAHOOTS expansions began in Eugene or Springfield and the additional hours of day each expansion covered.

Table A3: Reduced form effects of CAHOOTS expansions

	Call responses					Call outcomes	
	(1) Any response	(2) CAHOOTS response	(3) Only CAHOOTS	(4) Only police	(5) Both	(6) Arrest	(7) Use of force
Eugene 3-7 a.m., after Jan. 1 2017	-0.0289	0.116***	0.0602***	-0.145***	0.0560***	-0.0449*	0.000216
	(0.0293)	(0.0113)	(0.00743)	(0.0301)	(0.00898)	(0.0251)	(0.00670)
Eugene 7-10 a.m., after Nov. 1 2016	-0.0299	0.0848***	0.0480***	-0.115***	0.0368***	-0.0290	-0.00320
	(0.0305)	(0.0123)	(0.00801)	(0.0312)	(0.00966)	(0.0213)	(0.00564)
Springfield noon-11 p.m., after Jan. 14 2015	0.0214**	0.0638***	0.0270***	-0.0425***	0.0368***	-0.00507	0.000485
	(0.00884)	(0.00318)	(0.00240)	(0.00898)	(0.00219)	(0.00622)	(0.00169)
Springfield 9-11 a.m. and 11 p.m.-2 a.m., after Apr. 1 2017	-0.00139	0.0476***	0.00764**	-0.0490***	0.0400***	-0.00959	0.000354
	(0.00985)	(0.00482)	(0.00348)	(0.0104)	(0.00348)	(0.00765)	(0.00236)
Springfield 3-9 a.m., after Sept. 1 2017	-0.0196	0.111***	0.0500***	-0.131***	0.0611***	-0.0320**	-0.000112
	(0.0201)	(0.0110)	(0.00784)	(0.0210)	(0.00812)	(0.0146)	(0.00394)
Variance-weighted avg. effect	-0.020	0.094	0.045	-0.114	0.049	-0.031	-0.001
	(.01579)	(.00691)	(.00454)	(.01619)	(.00545)	(.01255)	(.00328)
Avg. effect	-0.012	0.085	0.039	-0.096	0.046	-0.024	-0.000
	(.01169)	(.00522)	(.00349)	(.01204)	(.00405)	(.00907)	(.00242)
Observations	284,487	284,487	284,487	284,487	284,487	284,487	284,487

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

Table A4: Call characteristics and outcomes by response

	(1) No response	(2) CAHOOTS only	(3) Police only	(4) Both
Panel A: Call type				
Welfare check	0.040	0.416	0.038	0.136
Suicidal person	0.007	0.087	0.011	0.149
Traffic/driver issue	0.175	0.052	0.084	0.035
Disorderly subject	0.048	0.063	0.092	0.156
Intoxicated person	0.003	0.042	0.003	0.012
Suspicous person	0.047	0.017	0.084	0.046
Subject down	0.006	0.014	0.002	0.007
Overdose	0.016	0.006	0.009	0.027
Mischief/trespassing	0.049	0.006	0.058	0.035
Dispute	0.051	0.004	0.253	0.155
Assault	0.007	0.001	0.031	0.028
Harassment	0.011	0.001	0.020	0.007
Burglary/theft	0.048	0.001	0.047	0.011
Drunk driver	0.065	0.000	0.024	0.003
Shots fired	0.014	0.000	0.015	0.001
Armed subject	0.001	0.000	0.008	0.004
Fight	0.002	0.000	0.010	0.002
Transportation	0.003	0.041	0.001	0.006
Assist police department	0.013	0.178	0.002	0.033
Mental issue	0.003	0.016	0.010	0.015
Assist fire department	0.005	0.044	0.006	0.011
Cardiac arrest	0.011	0.001	0.006	0.018
Robbery	0.001	0.000	0.005	0.002
Warrant	0.000	0.000	0.009	0.005
Incomplete call	0.110	0.001	0.024	0.006
Other	0.266	0.022	0.150	0.094
Panel B: Predicted call responses and outcomes				
Predicted CAHOOTS response	0.066	0.414	0.068	0.217
Predicted police response	0.516	0.443	0.678	0.636
Predicted arrest	0.069	0.057	0.103	0.094
Predicted use of force	0.003	0.003	0.006	0.005
Panel C: Call outcomes				
Time until dispatch (minutes)	5.668	31.003	13.227	10.865
Time until arrival (minutes)	0.241	39.031	18.474	17.434
Arrest	0.001	0.000	0.144	0.141
Use of force	0.000	0.000	0.007	0.011
Observations	89,307	13,584	148,721	10,541
Total	262,153			

Table A5: Traditional diff-in-diff and Borusyak, Javier, and Spiess (2024)

	CAHOOTS	Police	CAHOOTS and no police	Arrest	Force
	(1)	(2)	(3)	(4)	(5)
Panel A: Traditional diff-in-diff					
Post expansion	0.071***	-0.026***	0.028***	-0.014***	0.000
	(0.003)	(0.007)	(0.002)	(0.005)	(0.001)
Dependent mean	0.090	0.593	0.052	0.084	0.004
Observations	284,487	284,487	284,487	284,487	284,487
Panel B: BJS imputation estimator					
Post expansion	0.064***	-0.038***	0.025***	-0.023**	-0.004
	(0.003)	(0.015)	(0.002)	(0.011)	(0.003)
Dependent mean	0.006	0.710	0.003	0.116	0.007
Observations	31,980	31,980	31,980	31,980	31,980

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-city and call month FEs. To better 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: First stage of dispatch type on call busyness

	(1) CAHOOTS	(2) Police
Open CAHOOTS calls (SD)	-0.0207***	0.00379**
	(0.00104)	(0.00168)
Open police calls (SD)	-0.00130	-0.0544***
	(0.00143)	(0.00234)
Off-diagonal elements (p)		0.051
Observations	230,995	230,995

Notes: This table reports coefficients from a stacked regression of dispatch type (police or CAHOOTS) on open CAHOOTS and police calls. Call month by day of week by hour of day by city fixed effects are included. Robust standard errors, clustered by call are included in parentheses. The p-value is from a joint test of the off-diagonal coefficients being 0.

Table A7: Composition of compliers for each design

	CAHOOTS weights		Police weights
	(1) Diff-in-diff	(2) Busyness	(3) Busyness
Police only → both (π_{pb})	0.596	0.303	
	(0.023)	(0.027)	
Police only → CAHOOTS only (π_{pc})	.359	.183	
	(0.100)	(0.081)	
No response → CAHOOTS only (π_{pb})	0.045	0.514	
	(0.098)	(0.077)	
CAHOOTS only → both (π_{cb})			0.130
			(0.017)
CAHOOTS only → police only (π_{cp})			-.0238
			(0.026)
No response → police only (π_{np})			0.894
			(0.020)

Table shows shares for each of designs and compliance groups.

Table A8: Complier group-specific treatment effects on arrest

	Diff-in-diff		+ busyness	
	(1)	(2)	(3)	(4)
Police → both	-.171	-.114	-.159	-.0651
	(.243)	(.147)	(.168)	(.136)
Police → CAHOOTS	-.286	-.369	-.304	-.352
	(.345)	(.194)	(.199)	(.193)
None → CAHOOTS	.13		.102	
	(.447)		(.107)	
J test of overidentification	.0871	.171	.0913	.997
p	.957	.982	.993	.91

Each estimate inverse variance-weighted.