

Deterrence or Backlash?

Arrest and the Dynamics of Domestic Violence

Sofia Amaral¹ Gordon Dahl² Victoria Endl-Geyer³
Timo Hener⁴ Helmut Rainer¹

¹ifo Institute, LMU, CESifo ²UCSD, CESifo
³ifo Institute, LMU ⁴Aarhus University, CESifo

November 2022

*The views expressed are solely our own
and do not necessarily represent the opinions of our partners.*

Motivation

- **Domestic violence** (DV) is a serious and pervasive threat to the well-being of women worldwide
 - ▶ One third of women report some form of physical or sexual violence from a partner in their lifetime (WHO Report, 2021)
- A key aspect of DV is its **repeat nature**: women typically experience multiple offenses by the same partner
 - ▶ In our setting: Almost half of victims have a repeat call within 12 months
- How should police officers respond to domestic violence incidents in order to stop the cycle of DV?
- A highly controversial police response is to **arrest** suspects on the spot

Deterrence or Backlash?

① Proponents argue that arrest:

- ▶ induces a short-term incapacitation effect
- ▶ signals a high cost of repeat offenses in the future (deterrence)
- ▶ empowers women since they know something is being done

② Opponents argue that arrest:

- ▶ comes with a weak dosage of punishment
- ▶ may cause an escalation of DV (backlash)
- ▶ disproportionately impacts marginalized individuals, harming both them and their communities

For an overview of the arguments, see Goodmark (2017) "Should Domestic Violence Be Decriminalized?" Harvard Journal of Law and Gender

Research Question and Challenges

Does arrest increase or decrease repeat domestic violence?

Challenge #1:

- **Data limitations:** Measures of repeat victimization are largely not available in the context of DV
 - ▶ DV is usually not even a crime category, but must be inferred
 - ▶ Researchers typically only observe victimization conditional on a criminal charge (and DV charges often not filed)
 - ▶ Hard to identify & follow victims over time

Solution #1:

- **Unique administrative data:** Universe of all 999 emergency calls recorded by a major police force in the UK over 10 years
 - ① We observe the initial classification of DV by the call handler
 - ② We can match all DV calls to the arrest decisions made by officers
 - ③ We observe the geolocation of the dispatch for each call

Research Question and Challenges

Does arrest increase or decrease repeat domestic violence?

Challenge #2:

- **Selection bias:** Arrest is endogenous
 - ▶ Cases which result in arrest are more serious, and hence likely to be positively related to the underlying risk of repeat violence

Solution #2:

- **Identification:**
 - 1 Conditionally random assignment of police officers to 999 DV calls
 - 2 Heterogeneity in officers' propensity to arrest

Arrest is effective at breaking the cycle of repeat DV

- Arrest reduces repeat DV calls within the next 12 months by 49 pp
 - ▶ Amounts to a 51% reduction relative to the control complier mean
- The drop in repeat DV calls is a **reduction in incidence**, not a **drop in reporting**
 - ▶ Test 1: Severity of repeat DV calls
 - ▶ Test 2: Victim versus third-party reports of repeat victimization
- Mechanisms
 - ▶ **Short-term “cooling off”**: 23% of control compliers experience repeat DV within 48 hours; arrest prevents almost all of this revictimization
 - ▶ **Longer-term deterrence**: Additional reductions over the following year: \approx 50% reduction
 - ▶ **Criminal sanctions**: Rise in formal criminal charges

Prior literature

- 1 1981 Minneapolis DV Experiment and its 5 replications
 - ▶ Arguably doesn't answer the right question
 - ▶ RCTs with low rates of compliance and contamination issues → Mixed evidence
 - ▶ Use of an RCT challenging for anything other than ITT
 - ▶ Experiments moving beyond ITT would likely not get IRB approval
- 2 No-drop policies and mandatory arrest laws
 - ▶ Chin and Cunningham 2019; Aizer and Dal Bo 2009; Iyengar 2009
 - ▶ State-level analyses; ITT effects of laws
- 3 Judge stringency IV
 - ▶ Aizer and Doyle 2015; Kling 2006; Bhuller et al. 2018, 2020; Dobbie et al. 2018; Di Tella and Schargrodsky 2013; Muller-Smith 2021
- 4 Determinants of DV
 - ▶ Aizer 2016; Anderberg et al. 2016; Hidrobo et al. 2016; Tur-Prats 2021; Guarnieri and Rainer 2021; Bhalotra et al. 2021; Green et al. 2021; Card and Dahl 2011; Angelucci 2008; Stevenson and Wolfers 2006

Minneapolis DV Experiment

Table One
Designed and Delivered Police Treatments in Domestic Assault Cases

Designed Treatment	Delivered Treatment			Total
	Arrest	Advise	Separate	
ARREST	98.9% N=91	0.0% N=0	1.1% N=1	29.3% N=92
ADVISE	17.6% N=19	77.8% N=84	4.6% N=5	34.4% N=108
SEPARATE	22.8% N=26	4.4% N=5	72.8% N=83	36.3% N=114
TOTAL	43.3% N=136	28.3% N=89	28.3% N=89	100% N=314

Context

The West Midlands

- Second most populous county in England (≈ 3 million people)
- Includes Birmingham, Coventry, Wolverhampton
- Ethnically diverse
 - ▶ White 79%, Asian 11%, Black 3%, Other 7%
- Estimated annual IPV rate
 - ▶ WM = 6.6%, England = 5.8%
- Most DV occurs at home in England (90%)

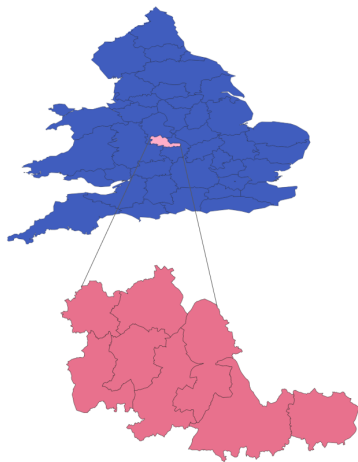


Figure: WMPF jurisdiction

Domestic Violence 999 Emergency Calls in West Midlands

- 3,448 emergency calls of DV per month
 - ▶ → 6.2 calls per 10,000 pop
- 9% of all calls are DV
- Compulsory to dispatch police officers when there is a DV call
- There is no specialization of first-response officers by type of offense

The Police Response System in the WMPF

- 1 **Call Handling:** Call handler records
 - ▶ Incident type (DV 9%, non-DV 91%)
 - ▶ Caller identity (victim 33%, third party 67%)
 - ▶ Priority level (1: immediate 60%, 2: early 20%, 3+: other 20%)
 - ▶ GPS location of incident
- 2 **Dispatch:** Dispatcher coordinates first response teams in the field
 - ▶ Monitors availability of patrol cars on electronic map
 - ▶ Assigns patrol cars to incidents [Example](#)
- 3 **First Response:** First response officers take one of four actions
 - ▶ (a) Arrest (3.1%), (b) Recommend criminal investigation (42.0%), (c) Provide advice/warning (9.3%), (d) No further action (45.6%)
- 4 **Investigation:** Investigator reviews case and decides whether to
 - ▶ Open criminal investigation (59.1%)
 - ▶ Coordinate a charge with prosecutors (1.4%)

Identification

Identification

Omitted variable bias:

- Arrests and recidivism likely to both be positively correlated with unobserved factors related to severity → upward bias in OLS

Solution:

- Officers (and by extension, officer teams) have different propensities to arrest
- Assignment of officer teams is “as good as random” after conditioning on geographic area, time, and call grade
 - ▶ Emergency calls follow a Poisson process (Maxfield, 1982)
 - ▶ Difficult to predict exactly when and where demands on police resources will emerge
 - ▶ We condition on fixed effects for wards (subsets of catchment areas), year, month, day of week, hour, and call grade

Construction of the Instrument

Definition

For each call i , the IV is the weighted average arrest rate of dispatched team members (officers) to DV calls, excluding the current address

- Each officer is assigned an arrest rate based on DV calls they have been a part of
- Weights are the number of DV calls an officer has been a part of
- Use all past and future DV calls
- Only use officers dispatched **before the first officer arrives**
- Require at least 400 combined DV cases for a team

Conditional Randomness

Our design exploits the way 999 calls are handled:

- Call handlers determine priority level and classify broad crime type
 - ▶ DV singled out as a crime type
- Based on this information, dispatchers assign officer teams
 - ▶ Decisions are made under time-pressure
 - ▶ Allocate cars to calls based on grade and proximity between a police vehicle and an incident
- We observe similar information about the call as the dispatcher
 - ▶ Call grade, time (year, calendar month, day-of-week, hour-of-day) and incident location
- **Conditioning on the information set, we have random assignment**

Empirical Strategy

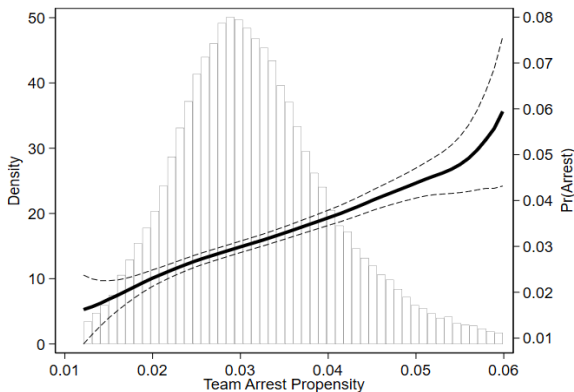
$$DV_i = \beta_0 + \beta_1 A_i + X_i' \delta + \epsilon_i \quad (1)$$

$$A_i = \alpha_0 + \alpha_1 S_{-i} + X_i' \theta + \epsilon_i \quad (2)$$

- X are the conditioning variables about the call observed by the dispatcher
- Standard errors clustered at the level of the officer with the most DV cases on the response team

First Stage

Figure: First stage of Arrest on Officer Stringency



Notes: This Figure displays the first stage of the baseline sample. The probability of an on the scene arrest is plotted on the right y-axis against officer stringency shown along the x-axis. Plotted values are mean-standardized residuals from regressions on call grade, ward, year, month, day, hour fixed effects. The solid line shows a local linear regression of arrest on officer stringency. The dashed line depicts 95% confidence intervals. The histogram shows the density of officer stringency along the left y-axis (top and bottom 1% excluded).

First Stage Coefficient

The first stage coefficient does not need to equal one for several reasons:

- The sample of cases used to calculate the stringency measure is not the same as the estimation sample
- There are covariates: fixed effects for wards (geographic areas), time (year, calendar month, day-of-week, hour-of-day), and call grade
- Officer teams are not held constant, with different officers belonging to different teams at different points in time

Data and Main Variables

Estimation sample	
Domestic violence cases classified by call handlers (2011-2016)	184,468
Nonmissing dispatch time	174,130
At least 400 DV cases in dispatched team	136,649
Call grade is 1 or 2 (baseline estimation sample)	124,216

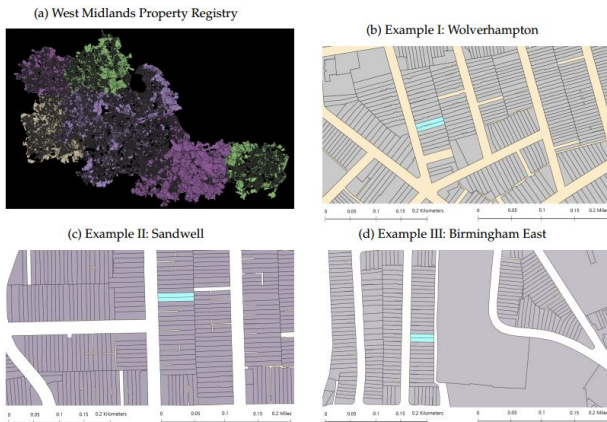
Instrument construction sample	
Officer-case level observations in call handler defined DV cases (2010-2019)	631,834

Measurement of Repeat DV

- Measure repeat DV calls against the same victim within 12 months
- Repeat DV identified via GPS location (10x10m level)
- 99.7% of DV calls have a geolocation

Measurement of Repeat DV, cont.

Figure 4: Property size in the West Midlands



Notes: In Figure (a) we display the polygons of registered freehold properties in the West Midlands division. In Figure (b)-(d) we display examples of residential areas in Local Police Units of Wolverhampton, Sandwell and Birmingham East, respectively. In these areas, the average property size are 157 sq. meters (12 x 12m), 145 sq. meters (12 x 12 m) and 199 sq. meters (14 X 14m), respectively. Source: Her Majesty's Land Registry INSPIRE Index Polygons.

- Lot sizes are between 12X12m -19X19m

Results

Balancing

	Dependent variable:	
	Arrest x 100	Team arrest propensity x 100
	(1)	(2)
Past DV history:		
Case in past 12 months	0.484** (0.172)	-0.015 (0.011)
Arrest in past 12 months	1.777*** (0.325)	0.016 (0.014)
Formal investigation in past 12 months	0.163 (0.184)	0.015 (0.011)
Criminal charge in past 12 months	2.043*** (0.305)	-0.014 (0.013)
Case characteristics:		
Caller identity (=1 victim)	0.161 (0.110)	0.013 (0.008)
Gender of call handler (=1 female)	-0.126 (0.126)	0.006 (0.007)
Call handler experience (years)	0.008 (0.008)	0.000 (0.001)
Mean of dep. var.	3.120	3.166
Joint F-statistic	22.936	0.946
[p-value]	[0.000]	[0.478]
Observations	124,216	124,216

Dependent variable: **Arrest**

	Prior DV call		DV hotspot		Time of day	
	Yes	No	Yes	No	Day	Night
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Baseline instrument						
Team arrest propensity	0.656*** (0.103)	0.799*** (0.083)	0.799*** (0.090)	0.656*** (0.085)	0.543*** (0.097)	0.853*** (0.083)
Mean of dep. var.	0.037	0.026	0.030	0.032	0.024	0.037
Observations	58,139	66,049	52,922	71,294	53,164	71,036
Panel B: Reverse sample instrument						
Reverse team arrest propensity	0.570*** (0.105)	0.674*** (0.082)	0.682*** (0.088)	0.441*** (0.083)	0.424*** (0.083)	0.319*** (0.078)
Mean of dep. var.	0.038	0.027	0.031	0.032	0.024	0.037
Observations	52,865	59,680	47,554	65,640	48,907	65,697

Main Result

	Dependent variable: Repeat call for DV			
	OLS	IV		
	(1)	(2)	(3)	(4)
Arrest	0.001 (0.008)	-0.517*** (0.170)	-0.488*** (0.187)	-0.488*** (0.187)
Call grade, time, ward F.E.'s	yes	yes	yes	yes
Ward x time F.E.'s	yes	no	yes	yes
Ward x call grade F.E.'s	yes	no	no	yes
Mean of dep. var.		0.492		
Control complier mean		0.962		
First stage		0.772*** (0.068)	0.723*** (0.070)	0.722*** (0.070)
Reduced Form		-0.400*** (0.131)	-0.353*** (0.136)	-0.352*** (0.136)
Kleibergen-Paap Wald F statistic		128	108	107
Observations	124,216	124,216	124,216	124,216

Reporting vs. Incidence

Do arrests lead to a reduction in repeat incidence or merely changes in reporting behavior?

- Two possible changes in reporting behavior:
 - (i) Arrest encourages victims to report future incidents
empowerment/deterrence
 - (ii) Arrest discourages victims from reporting future incidents
intimidation/backlash

What do the literature & police say?

- The probability of reporting increases with the repeat nature/severity of the problem faced by a DV victim
- Victims more likely to report in future when police treat initial reports seriously

Test #1: Severity of Repeat DV Calls

- We write down a simple model of threshold reporting behavior
 - ▶ If arrest empowers a woman (belief that arrest deters future abuse)
→ she tolerates a lower level of abuse before calling 999
 - ▶ If arrest results in backlash (fear of reporting due to retaliation)
→ her threshold for reporting abuse rises
 - ▶ Which effect dominates in the population is an empirical question
- Can use severity of repeat calls to test this
 - ▶ Call handler grades the severity of each 999 call
 - ▶ Classify severe cases as call grade 1 (60%), remainder as less severe
 - ▶ **Decomposition:** Split total effect into "repeat and severe" versus "repeat and less severe"

Test #2: Who Reports Repeat Victimization?

- If a victim is intimidated by their partner after an arrest due to fears of retaliation (backlash), the likelihood of not reporting a repeat DV case should be markedly higher for victims than for third parties
- Rationale: Victims should be more worried about retaliation than neighbors
- **Decomposition:** Split total effect into "repeat reported by third party" versus "repeat reported by victim"

Tests for Reporting vs. Incidence

Dependent variable:

	Repeat call for DV	Low severity repeat call	High severity repeat call	Victim-initiated repeat call	Third party-initiated repeat call
	(1)	(2)	(3)	(4)	(5)
Arrest	-0.488*** (0.187)	0.064 (0.150)	-0.552*** (0.173)	-0.099 (0.157)	-0.390** (0.190)
Mean of dep. var.	0.492	0.192	0.300	0.182	0.311
Control complier mean	0.962	0.132	0.830	0.275	0.687
Observations	124,216	124,216	124,216	124,216	124,216

Mechanisms

Mechanisms

- ① Short-term “cooling off”
- ② Longer-term deterrence
- ③ Criminal sanctions

Short-term and Longer-Term Effects

Dependent variable: **Repeat call for DV in the specified time frame**

within 96 hours in hours 1-48 in hours 49-96 within 12 months (excl. hours 1-96) in months 1-6 (excl. hours 1-96) in months 6-12

	(1)	(2)	(3)	(4)	(5)	(6)
Arrest	-0.198* (0.104)	-0.197** (0.099)	-0.025 (0.055)	-0.450** (0.184)	-0.323* (0.185)	-0.244 (0.178)
Mean of dep. var.	0.069	0.054	0.018	0.469	0.364	0.273
Control complier mean	0.251	0.233	0.043	0.909	0.684	0.507
Observations	124,216	124,216	124,216	124,216	124,216	124,216

Criminal Sanctions

	Dependent variable: Criminal charge
	(1)
Arrest	0.104** (0.053)
Mean of dep. var.	0.014
Control complier mean	0.020
Observations	124,216

Robustness Checks

① Measurement error in repeat victimization

▶ Selective moving

- ★ Decrease happens within 4 months, so not much time to move
- ★ Survey evidence: 11 out of ~7,500 respondents say they moved due to a DV event
- ★ Observe actual victim IDs if an investigation – no evidence that arrest affects whether we can track them as repeat cases using GPS (10.6 versus 10.4 missing repeats, for arrest and no arrest, respectively)

▶ Multi-unit housing

- ★ About 80% of households in WM reside in single-family houses
- ★ Similar findings if we exclude areas with a high fraction of apartments, or the center of the city

② Exclusion restriction

③ Heterogeneous effects

④ Intensive margin, alternative samples and instruments

Robustness: Multi-Unit Housing

	Dependent Variable: Repeat call for DV	
	Only areas with min. 80% of HH's in detached houses	Excluding city center (3km radius)
	(1)	(2)
Arrest	-0.468* (0.263)	-0.441** (0.194)
Mean of dep. var.	0.479	0.491
Control complier mean	0.931	0.914
Observations	81,953	118,559

Exclusion Restriction I

	Dependent variable: Repeat call for DV					
	(1)	(2)	(3)	(4)	(5)	(6)
Arrest	-0.488*** (0.187)	-0.487*** (0.186)	-0.487*** (0.183)	-0.486*** (0.187)	-0.480*** (0.182)	-0.518*** (0.191)
Formal investigation		-0.049 (0.034)			-0.040 (0.035)	
Time on scene			0.002 (0.022)		0.009 (0.022)	
Recommend investigation				-0.044 (0.061)	-0.031 (0.061)	
Advice				-0.001 (0.053)	-0.000 (0.053)	
Instrument: Team arrest propensity	yes	yes	yes	yes	yes	yes
Instrument: FI propensity	no	yes	no	no	yes	no
Instrument: Time on scene propensity	no	no	yes	no	yes	no
Instrument: Recommend FI propensity	no	no	no	yes	yes	no
Instrument: Advice propensity	no	no	no	yes	yes	no
Control: Team characteristics	no	no	no	no	no	yes
Mean of dep. var.			0.492			
Control complier mean			0.962			
Observations	124,216	124,216	124,216	124,216	124,216	124,216

Exclusion Restriction II

	Dependent Variable: Repeat call for DV				
	(1)	(2)	(3)	(4)	(5)
Arrest	-0.488*** (0.187)	-0.511*** (0.187)	-0.487*** (0.183)	-0.561*** (0.190)	-0.551*** (0.188)
Instrumented: Team arrest propensity	yes	yes	yes	yes	yes
Control: FI propensity	no	yes	no	no	yes
Control: Time on scene propensity	no	no	yes	no	yes
Control: Recommend FI propensity	no	no	no	yes	yes
Control: Advice propensity	no	no	no	yes	yes
Mean of dep. var.			0.492		
Control complier mean			0.962		
Observations	124,216	124,216	124,216	124,216	124,216

Robustness

	Dependent Variable: Repeat call for DV				
	Intensive margin	IV: call grade 1 and 2	IV: Traditional I-o-o	Team DV cases: 300	Team DV cases: 500
	(1)	(2)	(3)	(4)	(5)
Arrest	-1.516* (0.813)	-0.514*** (0.198)	-0.474** (0.187)	-0.406** (0.184)	-0.441** (0.192)
Mean of dep. var.	1.337	0.493	0.492	0.491	0.494
Control complier mean	2.867	0.921	0.948	0.880	0.915
Observations	123,063	118,788	124,229	134,534	111,437

Heterogeneity

	Dependent Variable: Repeat call for DV			
	Formally investigated DV case in past 12 months	Pr(Arrest X) is high	At least one female officer on response team	Response team with above mean age
	(1)	(2)	(3)	(4)
Yes	-0.744** (0.308)	-0.626** (0.307)	-0.552** (0.242)	-0.322 (0.230)
No	-0.431** (0.185)	-0.513*** (0.190)	-0.474** (0.222)	-0.677*** (0.230)
Observations	124,216	124,216	124,216	124,211

Conclusion

- Long-standing debate by academics and policy-makers about how the police should best respond to domestic violence
- Very limited and conflicting evidence so far
 - ▶ Minnesota Domestic Violence Experiment
 - Flawed policy and experimental design + underpowered
 - ▶ Staggered intro of mandatory arrest laws in U
 - Mixed evidence of a deterrent effect; state-level ITT analysis
- We find evidence consistent with a large deterrence effect
 - ▶ Reduction in DV incidence, rather than a reduction in reporting

Thank you!

Dispatchers' screens

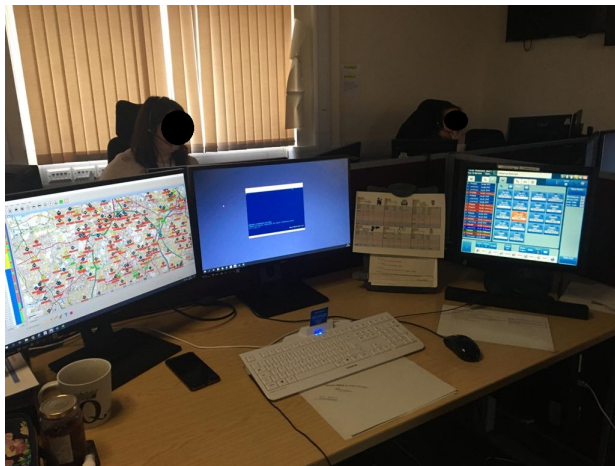


Figure: Visit to the dispatchers room in Sandwell, West Midlands - February 2020