Comparison of Police Agency Size and Killings by Police

Greg Stoddard   Dylan J. Fitzpatrick   John Greer

This repository contains code to recreate support analysis from our research paper "Predicting Police Misconduct". In our paper, we study how well a data-driven machine learning system can predict serious instances of police misconduct. The goal of such a system, often called an early intervention system, is to identify officers who exhibit warning signs of a serious negative event and intervene before with support and services before that comes to pass.

Analysis

Although our paper focuses on data from two large police departments - the Chicago Police Department and the New York Police Department - our further analysis suggests that police misconduct is not just a problem for large departments. Using data from the Mapping Police Violence project on people killed by police, we find 62% of police killings involve departments with fewer than 500 officers, and 47% of killings involve departments with fewer than 200 officers.

	rost_size_cat	cumulative_share_officers	cumulative_share_killings
0	1-99	0.35	0.33
1	100-199	0.47	0.47
2	200-499	0.60	0.62
3	500-999	0.69	0.74
4	1000+	1.00	1.00

Despite the fact that large and small departments account for a roughly proportionate share of police killings, smaller departments are much less likely to use a tool like an early intervention system. As of 2016, only 12.4% of departments used an EIS. By contrast, 80.5% of large police departments (500+ sworn officers), had implemented an EIS in 2016.

One roadblock for smaller departments to implementing a data-driven approach to early intervention is the resources necessary to construct such a system. Fortunately, in our research paper, we find evidence that smaller police departments might be able to get the benefits of an EIS using a simple policy like flagging officers with the most prior complaints in the past two years at a dramatically lower cost than building a full-blown machine learning system. As a caveat to the above analysis, we note that there are important differences between 'police misconduct' and police killings. Most instances of police misconduct, even serious misconduct, do not result in death. Similarly, not all police killings are the result of misconduct. With these caveats in mind, we used the MPV data because we otherwise lack national data on citizen complaints, use of force, or lawsuits.
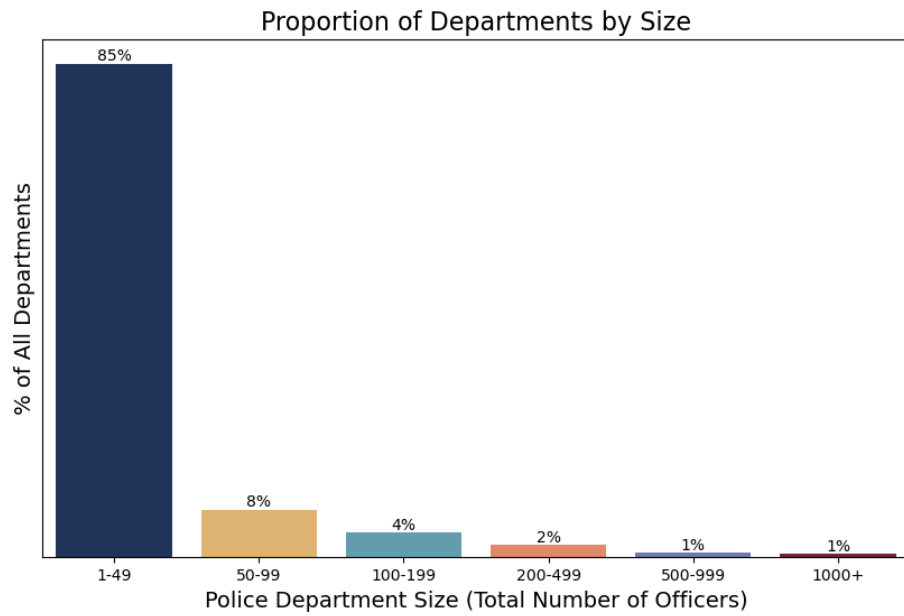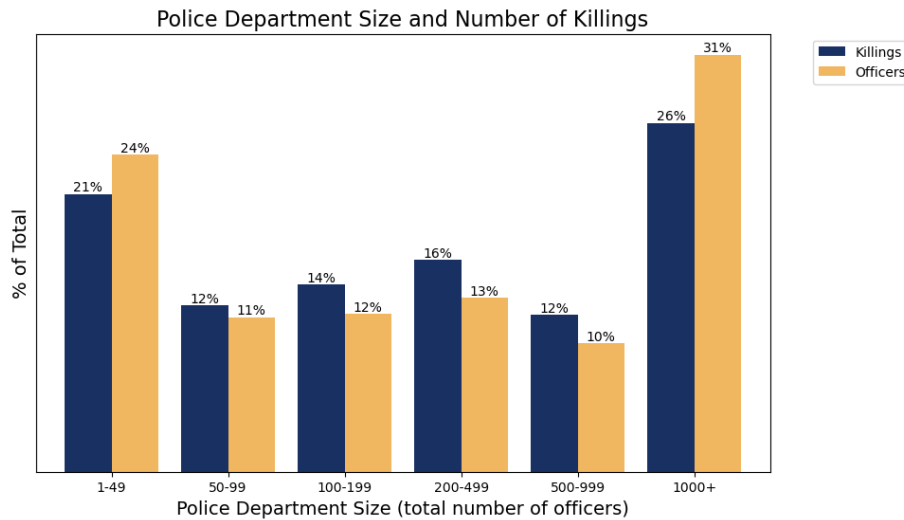
Figure 1: Histogram of agency sizes

Figure 2: Distributions of Police Agency Sizes and Proportion of Total Killings

Data & Methodology

We merged data from the Mapping Police Violence project (https://mappingpoliceviolence.org/) with the 2016 Law Enforcement Management and Administrative Statistics (LEMAS) survey to estimate the share of police killings that are committed by small versus mid-size versus large police departments. MPV constructs this dataset by monitoring a stream of news articles produced by Google News and then hand-verifying the details of a possible police killing. We merge this data to the LEMAS survey by agency name. In the cases where the MPV data lists multiple agencies as responsible, we attribute that event to each listed agency (those events are rare and have no qualitative impact on these findings). We limit our study to MPV-collected killings that occurred between January 2013 and April 2023.

Citations

"Mapping Police Violence," Campaign Zero, June 22, 2023, https://mappingpoliceviolence.org/

United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. Law Enforcement Agency Roster (LEAR), 2016. Inter-university Consortium for Political and Social Research [distributor], 2017-04-05. https://doi.org/10.3886/ICPSR36697.v1