Measuring Public Sentiment towards the Police

Emma Nechamkin

April 15, 2019

1 Introduction

Public perception of the police is incredibly important to police effectiveness and legitimacy but extremely difficult to measure. Public perception offers insight into how well a police department is functioning and may suggest adherence to tenets of procedural justice. Yet, compared to traditional performance metrics, metrics to evaluate public opinion are poorly defined and documented.

1.1 Clearance rates

Homicide clearance rates, or what share of murders a police department "solves", are a key performance metric for police departments. Chicago has one of the lowest homicide clearance rates in the country, and only about 1 in 6 murders lead to arrest. Moreover, Chicago's clearance rate has steadily declined over the past ten years, from about 40% in 2000 down to under 20% in 2017. As a comparison, several police departments have markedly higher clearance rates. Over the past decade, Los Angeles has solved 51% of murders and New York has solved 61% of murders.

    There are several potential reasons for the low clearance rate in Chicago, some of which suggest that non-traditional metrics of policing like procedural justice or public opinion may be related to traditional metrics. Police officers tend to cite the historically fraught relationship between the people and police, believing that someone who already views the police negatively because police seem inept may be less likely to cooperate with an investigation; more bluntly, many police officers lament a "no snitch" policy among victimized communities in Chicago. Evidence is conflicted: the National Crime Victimization Survey reports that these communities are no less likely to report crimes to the police, but a Cato Institute survey shows a race and education gap for crime reporting. There are also other viable explanations for Chicago's abysmal clearance rate, most notably that Chicago's police force has limited manpower per murder. Chicago has more murders than New York and Los Angeles combined, yet the police department (12,000 officers) is dwarfed by New York's (36,000) and Los Angeles' (10,000).

1.2 Public opinion

Procedural justice, or how police officers enforce laws, is necessary for effective policing. A civilian who considers the law enforcement process fair and just is likely to consider any related consequences fair and just, too. Conversely, when civilians perceive lack of procedural justice, they are more likely to file complaints and view their police force as delegitimate. For example, one study of New York Police Department Stop, Question, and Frisk stops showed that civilians who believed their stop to be fair were less likely to file a complaint than those who believed their stop

was unjust. Finally, a lack of procedural justice in just a few encouters can severely curtail public opinion of the police. Negative interactions with the police shape citizen perception up to fourteen times more strongly than positive ones.

Public perception of the police offers an additional metric to assess police performance. While hard metrics like clearance rates are easy to measure, assessing how the public feels towards the police is far more complex. Indeed, most work that tries to assess public sentiment uses survey-based or experimental research. Indeed, most past research has evaluated procedural justice through the lens of public opinion survey data. Such research is necessarily removed from the real world.

1.3   Research goals

Although public perception of the police is complex, I'm interested in assessing whether public perception of the police in related tweets is correlated with police effectiveness as measured by clearance rates. This follows recent work that has considered sentiments of tweets to assess public opinion of the police.

More specifically, I'm interested in assessing, in order of importance:

1. The extent to which public sentiment reflects traditional metrics of police effectiveness,
2. The differences by topic between effective and less effective police departments, and
3. How tweets about police departments reflect tenets of procedural justice.

As a caveat, public perception of the police is complicated and interacts with policing in myriad ways.

2   Past Work

2.1   Using twitter data to measure public sentiment towards the police

Although there has been limited work using data science techniques to study criminal justice, the Urban Institute used sentiment analysis for police-related tweets to measure how perception of the police changed due to the murder of Freddie Gray, using the following methods:

- Obtaining the data: researchers used a set of relevant tweets from 2014 and 2015 acquired through twitter.
- Processing the data: researchers removed mentions, hashtags, links, punctuation, and stop words from all tweets. They also used CoreNLP to tag tweets (e.g., to identify whether "cop" was a noun or a verb in each tweet).
- Learning models: researchers classified over 4,000 tweets manually to identify whether the tweet was positive, negative, neutral, or not applicable to their research for use in training and validation sets. They then used several types of models to predict the sentiment of new tweets and selected a gradient-boosted regression classifier as their model based on its accuracy (63%).
- Conclusions: researchers then used their newly labeled set of all tweets to assess the shift in public sentiment over time.

2.2 Using twitter data to connect public opinion with tweet sentiment

As a more general example, researchers at Carnegie Mellon University determined that public opinion surveys correlate to twitter sentiment on several key issues. They used twitter data specifically with two endgoals: to identify relevant tweets and to estimate sentiment (positive and negative) about a given topic.

In their work, researchers obtained tweets from 2008 and 2009 using the twitter API. They then used key words (like "obama" to measure presidential approval) to ensure that their tweets were relevant. Tweets were classified as positive, negative, or both depending on whether there was a positive, a negative, or both types of words in it. Finally, to get an accurate measure of sentiment, they computed a moving average aggregate of sentiment ratios, where sentiment ratio was defined as the ratio between the number of positive and negative relevant tweets. The moving averages allowed them to smooth otherwise volatile data. They then investigated correlations between the sentiment they uncovered and traditional public opinion surveys.

3 Logistics

3.1 Steps

For this project, I will specifically focus on Chicago, Los Angeles, and New York in several (broad stroke) steps.

- Obtain relevant data using the twitter API and/or python's tweepy package. Relevant tweets will contain hashtags related to Chicago police, New York police, and Los Angeles police
- Pre-process data by removing stop words and using baseline CoreNLP functionality and potentially python's NLTK package.
- Create a model to identify the sentiment of tweets (TBD based on results)
- Use topic modeling to evaluate how procedural justice is reflected in tweets, and what else is reflected in relevant tweets

3.2 Packages

I plan to use an AWS server to complete the CoreNLP work, and a variety of ML packages in python to build and evaluate models (include scikit-learn, and, depending on the models used, other more advanced packages). I also plan to use python's NLTK for topic modeling and tweepy/twitter's API.

3.3 Milestones

At minimum, I hope to have completed the first two steps listed (obtaining relevant tweets and preprocessing tweets using CoreNLP and manual tagging) by the mid-quarter presentation.

3.4 Project evaluation

By the end of the project, I hope to evaluate the sentiment using a testing / training / validation framework in which I assess how well my model has classified tweets.

Topic modeling is currently an aspirational part of this project and I expect to have a more concrete plan as we gain depth in this class.

3.5 Logistics

I am working alone and so do not need to coordinate work among teammates. My rough timeline is shown below. This timeline is aspirational and reflects the "best case" of work.

- April 19: Twitter data collected
- April 23: Basic preprocessing completed (CoreNLP tagging, similar to Urban Institute)
- April 26: Manual coding of tweet sentiment completed
- April 31: First run of selections of models completed
- May 3: Models evaluated
- May 10: Model refinement and adjustment
- May 15: Topic modeling outlined
- May 20: Topic modeling completed
- May 25: Analysis completed
- May 30: Report and presentation draft completed

4 Selected sources (informal)

4.1 Papers

- Ekins, Emily. (2016). Policing in America: Understanding Public Attitudes Toward the Police. Results from a National Survey. SSRN Electronic Journal. 10.2139/ssrn.2919449.
- Fowler, AF Rengifo and K. 2016. "Stop, Question, and Complain: Citizen Grievances Against the NYPD and the Opacity of Police Stops Across New York City Precincts, 2007-2013." Journal of Urban Health (93 Suppl 1): 32-41.
- O'Connor, Brendan & Balasubramanyan, Ramnath & R. Routledge, Bryan & A. Smith, Noah. (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. International AAAI Conference on Weblogs and Social Media. 11.
- Oglesby-Neal, Ashlin, Tiry, Emily, and Kim, KiDeuk. "Public Perceptions of Police on Social Media." 19 February 2019. Urban Institute Research Brief.
- Skogan, Wesley G. 2006. "Asymmetry in the Impact of Encounters With Police." Policing and Society.
- Tyler, Tom R. 2004. "Enhancing Police Legitimacy." The Annals of the American Academy of Political and Social Science 593: 84-99.

4.2 News and articles (quick links)

- https://www.washingtonpost.com/graphics/2018/investigations/unsolved-homicide-database/?utm_term=.8da8a801878a&city=indianapolis]
- https://chicago.suntimes.com/news/murder-clearance-rate-in-chicago-hit-new-low-in-2017/
- https://www.theatlantic.com/ideas/archive/2018/05/quis-custodiet-ipsos-custodes/560324/
- https://datasmart.ash.harvard.edu/news/article/map-monday-unsolved-homicides