I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16609

Strategic Bureaucratic Opacity: Evidence from Death Investigation Laws and Police Killings

Elda Celislami
Stephen Kastoryano
Giovanni Mastrobuoni

NOVEMBER 2023

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16609

Strategic Bureaucratic Opacity: Evidence from Death Investigation Laws and Police Killings

Elda Celislami
University of Reading

Stephen Kastoryano
University of Reading and IZA

Giovanni Mastrobuoni
University of Turin, Collegio Carlo Alberto, CEPR and IZA

NOVEMBER 2023

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Strategic Bureaucratic Opacity: Evidence from Death Investigation Laws and Police Killings*

Police accountability is essential to uphold the social contract. Monitoring the monitors is, however, not without difficulty. This paper reveals how police departments exploit specific laws surrounding death investigations to facilitate the underreporting of police killings. Our results show that US counties in which law enforcement can certify the cause of death, including counties which appoint the sheriff as the lead death investigator, display 46% more underreported police killings than their comparable adjacent counties. Drawing on a novel adapted-LATE potential outcomes framework, we demonstrate that underreported police killings are most often reclassified as 'circumstances undetermined' homicides. We also show that law enforcement agencies in counties with permissive death certification laws withhold more homicide reports from the public. The main underreporting results are primarily driven by underreporting of White and Hispanic deaths in our analysis sample, with the effect on Hispanic people particularly pronounced along the US-Mexico border. We do not find that excess underreported killings are associated with more violence directed towards police. We do, however, note a nationwide positive correlation between the permissiveness of gun-laws and underreported police killings. In addition, we find more underreporting in counties which have both high per-capita Google searches for Black Lives Matter and which allow law enforcement to certify the cause of death. Our results do not indicate that other differences in death investigation systems - coroner vs. medical examiner, appointed vs. elected, or physician vs. non-physician - affect the underreporting of police killings.

JEL Classification: K42, H11, K13

Keywords: police killings, police violence, death investigations, coroner, medical examiner, policing

Corresponding author:
Stephen Kastoryano
Department of Economics
University of Reading
Shinfield Rd
Reading RG6 6EL
United Kingdom
E-mail: s.p.kastoryano@reading.ac.uk

* Thanks: Judy Melinek, Samuel Sinyangwe, Christian Dustmann. We also thank seminar speakers at the LSE Crime Seminar, King's College London, the Hertie School of Government, the University of Bergen 'Workshop on Health, Human Capital and Social Insurance,' the RWI Berlin.

1 Introduction

Any discussion of state power is a discussion of its limits. Ideally, state institutions would always operate within their legal mandates and in view of socially optimal outcomes. In reality, of course, this is not the case. Institutional mechanisms must, therefore, guard against potential misconduct and ensure accountability where it occurs. Ensuring accountability is particularly relevant when it comes to law enforcement, given their entrusted power, and perhaps most symbolic when it comes to police killings (Sherman, 1978). Failure to keep police accountable can affect trust in the state and have far-reaching social consequences as evidenced by the Black Lives Matter movement. Unfortunately, monitoring law enforcement and police killings is not without difficulty. Career-minded police officers and their superiors, veiled behind the 'blue wall of silence,' have an interest in avoiding lengthy investigations and any financial compensation to relatives of wrongfully killed victims (Zimring, 2017).

As it stands, relevant institutions fail at the very first step: accounting for the scale of the problem. Over the years 2013-2019, official records from the FBI's Supplementary Crime Reports only disclosed half of the total police killings documented in open-source registries. It remains unclear which failures of monitoring mechanisms allow such widespread opacity in reporting by local police departments – reporting that is not subject to any state or federal auditing – and whether this opacity is strategic.

Of particular concern are certain pressures and historical laws embedded in the death investigation system. The National Association of Medical Examiners has documented that about a fifth of coroners and medical examiners[1] had been put under pressure by public officials to change the cause or manner of death (Luzi et al., 2013).[2] In addition, 23 states include some counties in which law enforcement is legally permitted to directly certify the cause of death under certain conditions. A subset of these counties also supports systems with notably adverse incentives in which the sheriff is appointed as the de facto lead death investigator.

This study evaluates the causal effect of different laws and systems surrounding death investigations on the underreporting of police killings, implicitly testing for a conflict of interest in reporting. The study further offers empirical evidence of specific methods used by law enforcement to lower officially reported police killings by drawing on various data sources and a novel adapted-LATE potential outcome framework with more general ap-

[1]Besides law enforcement's ability to certify the cause of death, state and county death investigation systems also differ along other important characteristics. In general, death investigation systems are separated into two categories: medical examiners and coroners. Medical examiners are always appointed, while over two-thirds of coroners are elected officials. Medical examiners also require a medical degree in a vast majority of states and are often trained as forensic pathologists. The requirements for coroner duties vary more substantially. In some states, sufficient credentials to examine a deceased person and certify the cause of death are to be 18 years of age, hold a high school degree and follow a 40 hour course.

[2]As a result, the National Association of Medical Examiners petitions "whistleblower protection" whenever death investigators uncover abuse or criminal activity in officer-involved shootings (Melinek et al., 2013)

plicability. The paper also investigates the drivers, moderating factors, and repercussions of our measured underreporting effect.

The analysis uses a newly constructed dataset for the years 2013-2019 which includes county-specific information on death investigation systems. Our primary outcome variable in the dataset is the difference between two variables. The first is the 'true' count of police killings from Mapping Police Violence (MPV)[3], an open source registry that opened in the aftermath of the Black Lives Matter movement. The second is the official count reported by law enforcement agencies in the Supplementary Homicide Report (SHR) of the FBI's Uniform Crime Reporting (UCR) program. Our main result indicates that counties with laws permitting law enforcement to certify the cause of death display 46% more underreported police killings than their comparable adjacent counties. This effect is not solely attributable to counties where the sheriff is the lead investigator.

Historically, all death investigations until the late 1800s were conducted by the coroner, a colonial vestige of English common law. Over the following century, many states and isolated counties recognized the need for more scientific autopsies and introduced various laws requiring, among others, the presence or oversight of a medical professional in forensic work. This shifting tide came, for no apparent reason, to an almost full stop in the mid 1990s, leaving a geographical patchwork of death investigation systems and laws differing across state lines but also within states (Hanzlick, 2007).

Our empirical approach takes advantage of the resulting plausibly exogenous variation, rooted in movements to prioritize more scientific autopsies rather than reduce illegitimate interactions with police. It also exploits the fact that legislation is drafted centrally for each state, but applies to all counties. These developments, which we elaborate on, allow us to compare adjacent counties, some of which lie across state borders, with different laws and systems. Accounting for adjacency cluster fixed effects, we show that the adjacent counties in our analysis are comparable along population, urbanization, crime, income, racial-ethic distributions, and political voting behavior, which are known to be correlated to underlying views on police violence. Within this context, our causal effects of interest are long-run differences in the reporting of police killings between neighboring counties with comparable potential for permissive cultures of police violence and corruption.

Beyond uncovering differences in underreporting depending on death certification laws, we also explore different possible mechanisms through which sensitive homicide data can be hidden from the public eye. To begin with, we explore whether counties in which law enforcement is permitted to certify the cause of death are more likely to reclassify homicides in their reports. Our results indicate that counties with permissive certification laws report neither higher total homicides nor higher law enforcement homicides. They do, however, exhibit 35% more deaths categorized as 'circumstances undetermined,' the residual SHR homicide category, pointing towards the reclassification

[3]As discussed later, we do not require the 'true' count to be free from error, as long as on average the 'true' measure is unbiased or the bias is unrelated to the death investigation system.

of sensitive homicides.

We further investigate whether the underreporting of police killings is linked to strategic withholding of crime statistics from the public. To investigate this question, we draw from recently released data sharing information for all 24,620 law enforcement agencies in the US. Our results show that agency participation rates for the FBI's UCR program are 3.6 percentage points lower for counties in which law enforcement has authority to certify the cause of death compared to their controls. Looking at agency participation for the National Incident-Based Reporting System (NIBRS), a database that includes more detailed information pertaining to the circumstances of a homicide, this difference grows to 18.3 percentage points. We also find that this effect is reduced and not significant when looking exclusively at sheriff-coroner counties, arguably owing to sheriff-coroners' independence in directly reclassifying the cause of death.

Co-movements in the effects of underreporting, 'circumstances undetermined' homicides, and data withholding, are suggestive of cover-up mechanisms. They dispel alternative explanations for the measured underreporting effects. For example, one might hypothesize that crowd-sourced measures, such as media or civilian reports, more rigorously track police killings when sheriffs are tasked with certifying the cause of death. This would lead to higher underreporting in law enforcement certifying counties. Were this the case, we would not expect the observed increases in the number of 'circumstances undetermined' homicides reported to the FBI.

Looking at average changes across homicide categories may not, however, answer all policy relevant questions. For example, the average effects cannot reveal the share of counties which underreport police killings by re-classifying them into 'circumstances undetermined' homicides rather than using other cover-up methods. It is also unclear whether the stated average effects are driven by counties in which law enforcement agencies would only underreport police killings when facing lenient certification laws. They may also be attributable to counties in which law enforcement agencies would always underreport police killings, but would use different cover-up strategies depending on the permissiveness of certification laws. To answer these types of questions about which population substrata drive our average effects, we develop an adapted LATE potential outcomes framework in the spirit of Imbens and Angrist (1994) and with parallels to the discussion of supercompliers in Comey et al. (2022).

Our adapted potential outcome framework follows Comey et al. (2022) by assuming two monotonicity assumptions. However, instead of applying them on the treatment and outcome, we apply them on two binary outcomes of interest. The first applies to the underreporting of police killings, the second applies to declaring excess 'circumstances undetermined' homicides. Thereafter, because we are not studying a classical instrument-treatment-outcome causal chain, we depart from previous LATE frameworks by discarding the exclusion restriction, which is not credible in our setting. Instead, partly justified by evidence, we exclude two subtypes of law enforcement behavior using a cross-monotonicity restriction in order to achieve identification. Our identification strategy can find wider appeal in causal studies with a binary treatment and two binary outcomes, where researchers are interested in decomposing average effects on each outcome into effects for different cross-outcome subtypes.[4] Among other results, one of the main findings from the proposed framework is that nearly 90% of counties which would only underreport police killings when subject to lenient certification laws would also simultaneously increase reported 'circumstances undetermined' homicides. We further find that among counties which always underreport police killings, an equal share avoid scrutiny by always substituting these killings into the 'circumstances undetermined' category as those which always exploit other avenues. In terms of policy, these results suggest that joint investigations of the two outcomes would be better able to predict genuine strategic underreporting.

Throughout the results, we further discuss the drivers, moderating factors, and repercussions of our measured underreporting effect. We find evidence that victims were more often said to display signs of mental health issues in law enforcement certifying counties, as reported most often by law enforcement to media outlets. We also see some indication that the main effect is driven by higher true police killings, as opposed to lower reported homicides, and that strategically reported subcircumstances justifying police killings are different depending on certification laws. These results lend some credence to the hypothesis that years of more limited accountability can result in a culture of impunity among police for lethal engagement. In contrast, citizens in counties permitting law enforcement to certify the cause of death do not seem to engage in more violent and lethal assaults against police. Instead, looking at nationwide correlations, we only note that lenient gun laws are positively associated to the underreporting of police killings. Taken together, these results may underline that fear, rather than true threat, is more relevant in determining wrongful, and subsequently reclassified, police killings. We do not find any moderating effects of body-worn cameras, but this may be due to the data source which imprecisely measures their actual usage. In a last analysis on the question of monitoring, we consider whether awareness and concern for issues raised by the BLM movement, as proxied by 'Black Lives Matter' Google search trends by state and their year-to-year changes, influence the underreporting of police killings. We show that states with higher concern for the BLM movement display more underreporting of police killings, and law enforcement certifying counties in states with high searches for 'Black Lives Matter' display more underreporting of police killings.

In terms of racial-ethnic groups, we find that the main underreporting is primarily driven by homicides of White people and, in some of our specifications, by Hispanic people. We do not find any effects on Black people. The evidence of underreporting

[4] For example, we can consider the development context of a cash transfer which shows positive effects on the probability of investing in livestock and in children's education. One may investigate whether an increase in the probability of investing in livestock is mainly driven by households which also use the cash transfer to increase investments in their children's education (complements), or not (substitutes).

effects on Hispanic people and the absence thereof for Black people parallels findings in Fryer Jr (2019), which directly looks at fatal police shootings. As we will further discuss, the somewhat surprising absence of effect on Black people could also result from our specific adjacent county analysis sample, which, while representative of the whole country on almost all observed measures, underrepresents Black people relative to the US average. We further investigate the results on Hispanic people showing that counties on the US-Mexico border display higher underreporting of police killings for Hispanic people than their controls more inland, but display no differences for other racial-ethnic groups.

In a final step, the paper also evaluates whether other characteristics of death investigation systems are associated with differences in the underreporting of police killings. The results from separate adjacent county analyses do not indicate that distinctions in coroner vs. medical examiner systems, appointed vs. elected systems, or physician vs. non-physician systems, are associated with differences in the underreporting of police killings.

Our paper contributes to several strands of literature. First, it adds to the literature on incentive structures and police corruption (Becker and Stigler, 1974; Sherman, 1978; Baicker and Jacobson, 2007; Banerjee et al., 2021; Owens and Ba, 2021). More specifically, with regard to the question of how to monitor the monitors (Rahman, 2012; Cheng and Long, 2018; Long, 2019; Mastrorocco et al., 2020; Devi and Fryer Jr, 2020), our article emphasizes the potential conflict of interest for officers reporting police violence, underlining the social cost of poorly designed death investigation laws when it comes to monitoring police for their killings. We also add to the literature describing patterns of fatal police violence (Edwards et al., 2018; GBD et al., 2021; Schwartz and Jahn, 2020) by identifying which mechanisms enable law enforcement to avoid accountability. In particular, we explain the role of related, and sometimes co-opted, institutions in aiding cover-up cultures. These come on top of known internal mechanisms within police departments and unions (Rad et al., 2023). The closest study to ours is by Prados et al. (2022) which, in a comparison of Californian counties over 2000-2018, finds higher underreporting of police killings in sheriff-coroner counties.

The paper also contributes to an underdeveloped, but growing, literature on police killings, especially when compared to the literature on the death penalty, the only other criminal justice program that justifies killings. While fewer than 20 inmates are executed each year, the true number of police killings is close to 1,000. Yet, between 2000 and 2009 not a single legal study mentioned "police killings", "police use of deadly force" or "police shootings in their abstracts, against 589 citing the words "death penalty," "executions" or "death sentence" (Zimring, 2017).

The remainder of the paper proceeds as follows. In the next section, we provide some background on the current understanding of police violence and its consequences and on death investigations and their historical contexts. In Section 3 we describe our data. Section 4 outlines our empirical identification strategy and our estimating equation. Section 5 contains the results and also presents our new potential outcomes effect decomposition framework. Section 6 concludes.

2 Background

2.1 Underreporting of police killings

Patterns and trends in police killings have until recently been difficult to assess due to data reporting issues. Data collected by the FBI through law enforcement agencies and in the National Vital Statistics System (NVSS) through state vital registers are widely acknowledged to undercount the true number of deaths involving police in the US (GBD et al., 2021; Edwards et al., 2018, 2020; Nix et al., 2017; Fryer Jr, 2019).

The fragmented understanding of officer-involved fatality statistics came under focus in the mid-2010s following several high-profile police killings.[5] In response to these events and the ensuing Black Lives Matter protests, several news and non-governmental open-source outlets such as The Counted from *The Guardian*, the Police Shooting Database from *The Washington Post*, Fatal Encounters and Mapping Police Violence began collecting and publishing information about fatal police violence across the US.

The analysis of these new data sources offered evidence for previously suspected underlying trends. In general, Black, Hispanic, and Indigenous people are more likely to fall victim to fatal police violence, and the risk of fatality is particularly pronounced for men in their young adulthood. Unsurprisingly, officer-involved fatalities are highest in dense urban areas but, when considered as a share of total homicides, vary considerably across US census divisions. For example, in mountain states,[6] police were responsible for about 17% of all homicides between 2012-2018, while in mid-Atlantic states,[7] police accounted for 5% of all homicides. Officer-involved fatalities as a share of total homicides are also sizeable in smaller rural and non-core areas where police accounted for more than 10% of all homicides in contrast to 7% of all homicides in large central metropolitan areas. All in all, between 2012 and 2018, more than 1 in 12 of all homicides of adult men came at the hands of police (Edwards et al., 2018; GBD et al., 2021; Schwartz and Jahn, 2020).

Using secondary open source data on police killings, recent papers have explored which regional characteristics correlate with the underreporting of police killings. Feldman et al. (2017) and GBD et al. (2021) compare the NVSS, a dataset with similar fractions of underreporting as the SHR over our period of interest, to several open source datasets. They find that underreporting of police killings falls between 50% and 60% for Black,

[5]These included, but were not limited to, the shootings of Michael Brown, Trayvon Martin, Oscar Grant, Charleena Lyles, Stephon Clark, Eric Garner, Sandra Bland, Renee Davis, Philando Castille, Laquan McDonald, and Tamir Rice.

[6]Montana, Idaho, Wyoming, Nevada, Utah, Colorado, Arizona, and New Mexico

[7]New Jersey, New York, and Pennsylvania

White and Hispanic people, with less underreporting for other minorities.

Beyond the direct consequence of lost lives, there exist many other repercussions for a society facing high levels of police violence (Owens and Ba, 2021). The threat of potentially lethal police violence on unarmed citizens within a neighborhood has been found to directly increase mental health issues (Bor et al., 2018; Tyler et al., 2014) and negatively affect students' school attendance and schooling outcomes (Ang, 2021). These immediate effects have also been shown to corrode the implicit social contract between the state and citizens in the form of lower cooperation with law enforcement (Ang et al., 2021) and reduced willingness to report violent crime (Desmond et al., 2016; Zaiour and Mikdash, 2023). Notably, however, police killings, in particular of Black and Hispanic victims, have been shown to spur voter turnout (Ang and Tebes, 2020) and citizen support for civilian review boards and criminal justice reforms (Olzak, 2021; Morris and Shoub, 2023).

Several more general problems were also emphasized during the Black Lives Matter movement in relation to officer-involved fatalities. The first issue was that of implicit bias in policing. This bias is backed with much evidence in the case of street and vehicle searches (Gelman et al., 2007; Knowles et al., 2001; Antonovics and Knight, 2009; Pierson et al., 2020; Goncalves and Mello, 2021), and when it comes to white-officer interactions in minority neighborhoods (Hoekstra and Sloan, 2022). It has not, however, been decisively confirmed for fatal police shootings (Nix et al., 2017; Ross, 2015; Fryer Jr, 2019). A second highlighted problem was bias in the judicial processes and sentencing for minorities. While most empirical studies support these claims of bias, not all are in agreement on how to best address the problem of unobservable factors in sentencing (Arnold et al., 2018; Anwar et al., 2012; Rehavi and Starr, 2014; Abrams et al., 2012; Feigenberg and Miller, 2021).

A last problem emphasized was the prevalence of systematic and institutionalized permissive cultures of police violence which lack accountability. Proponents of reform argue that police are rarely held accountable for their misconduct. This lack of accountability is aided by formal union agreements that impede internal audits and the creation of civilian review boards (Rad et al., 2023), but also by the existing informal cover-up culture in police departments, known as the 'blue wall of silence' (Benoît and Dubra, 2004). Addressing culture is not without complications. Internal monitoring mechanisms must rely on officers breaking norms of police solidarity to investigate their own colleagues. Besides social shunning, these perceived betrayals can result in safety threats when officers require support (Rahman, 2012). External monitoring mechanisms, such as external review boards or added media scrutiny, can avoid these conflicts of interest but, by raising the expected penalty of an officer's errors, may in theory reduce the effectiveness of police (Prendergast, 2003; Gavazza and Lizzeri, 2007).[8]

[8] There exists some empirical evidence to support this view, showing that various periods of increased oversight on police led to de-policing and higher crime in the short run (Shi, 2008; Ba and Rivera, 2019; Devi and Fryer Jr, 2020; Premkumar, 2020; Campbell, 2022).

2.2 Death Investigation Systems in the US

Death investigators in the US play a key role in the detection of emerging diseases, in accounting for drug or consumer product related deaths, and in assessing violent deaths involving homicide and suicide (Hanzlick, 2007). These death investigation systems are divided into two main streams: the coroner system and the medical examiner system.

In general, medical examiners are appointed physicians, often specialized as forensic pathologists.[9] Coroners are most often elected officials without medical degrees or extensive forensic training (Hanzlick, 2007).[10] A common lower bar for eligibility as a coroners is to be older than 18-21 years of age, a county resident, and hold a high school degree. Coroners are also required to go through an initial 40 hour training session over their first year, with 17 of 28 states requiring an additional 8 hour training session every subsequent year (Ruiz et al. (2018), CDC training).

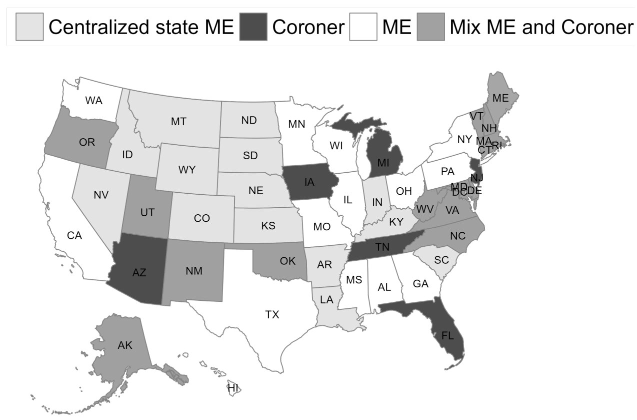


Figure 1: Death investigation systems in the US. Source: CDC

Historically, the medical examiner system evolved out of the coroner system, a relic of British colonialism, as outlined in the *Model Postmortem Examination Act* published in 1954. The stated purpose of the act was to promote systems capable of competently overseeing postmortem investigations, in particular for death cases where criminal liability may be involved (Hanzlick, 2014; Leflar, 1955). The call to address the medical incompetence of coroners was followed by a wave of states moving from coroner to medical examiner systems or mixed systems between the early 1960s and mid 1990s.[11] Since 1996, no additional state has moved away from a coroner system and county-specific transitions are rare (Hanzlick, 2007, 2014).[12]

At present, coroners and medical examiners are, depending on the state, distributed into one of four systems as shown in Figure 1. 16 states (and DC) have a medical examiner system in which death investigations are centralized in one location. 6 states

[9]Some exceptions apply. Wisconsin and West Virginia do not require medical examiners to be physicians (Hanzlick, 2007).

[10]Only 4 states, Kansas, Louisiana, Minnesota, and Ohio, require coroners to hold a medical degree. In addition, in Nebraska, the office of coroner is jointly held with the county attorney, so all coroners are lawyers, while 15 of 39 counties in the state of Washington have coroner duties performed by the county's prosecuting attorney.

[11]The chronology of these changes is presented in Hanzlick (2007).

[12]Hanzlick (2007) found that only 8 counties out of 3,143 shifted from a coroner system to a medical examiner system between 1996 and 2007.

have a decentralised county- or district-based medical examiner system. 14 states have a decentralised county-, district-, or parish-based coroner system, and 14 states have a decentralised county-based system with a mixture of coroner and medical examiner offices. 25 of the 28 states with coroners also have an appointed state medical examiner who, upon request, may perform autopsies and provide expert advice, but the coroner still ultimately establishes the cause of death (CDC; Fernandez, 2019). In 2019, of the 3,143 counties in the US, only 37% were served exclusively by a medical examiner but these cover 63% of the total population.[13]

Feldman et al. (2017) and GBD et al. (2021) allude to the question of death investigation systems, mentioning that they find no significant differences in underreporting depending on whether the county had a coroner or medical examiner system. In an analysis limited to California over 2000-2018, Prados et al. (2022) compare sheriff-coroner counties to all others and find higher underreporting of police killings in the former. None of these articles adopts causal analysis methods.

2.3 Certifying the Cause of Death

In general, coroners and medical examiners are placed in a vulnerable position with respect to law enforcement. The autopsies they perform must draw from background information and discussions with law enforcement, but the final assessments of the cause of death should, in principle, remain independent of external influence. In practice, this separation is not maintained. In 2013, the National Association of Medical Examiners conducted a survey on its members' independence in death investigations. It found that over 20% of forensic pathologists working for a medical examiner or coroner reported they were pressured by elected officials or appointees to change their findings. A quarter of those who resisted pressure - threats, termination, intimidation, legal action - suffered consequences for their actions. In addition, approximately 10% of respondents had been asked to sign death certificates or autopsy reports which were inconsistent with the findings of the autopsy (Melinek et al., 2013; Luzi et al., 2013).

The negative results of this conflict of interest are likely exacerbated in legal frameworks which grant stronger authority to law enforcement in the death certi-

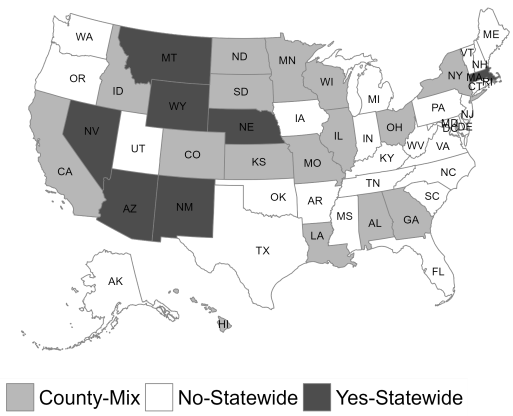


Figure 2: Law enforcement certification of cause of death.

[13]Numbers derived from our data.

fication process. Currently, 20
states include a subset of counties in which sheriffs hold the dual post of sheriff-coroner. Of these, some states (California, Nevada, North Dakota) designate the sheriff as the de-facto coroner under some conditions. In addition, Arizona, Massachusetts, Montana, Nebraska, Nevada, New Mexico, and Wyoming allow law enforcement to complete death certificates when the examiner is not present or is unavailable (Ruiz et al., 2018). As displayed in Figure 2, this results in 16 states in which a subset of counties allow law enforcement to certify the cause of death and 7 states in which all counties allow law enforcement to certify the cause of death. Allowing law enforcement to classify and/or certify the cause of death gives rise to concerning conflicts of interest when it comes to officer-involved fatalities. These conflicts of interest constitute the main focus of our analysis.

3 Data

3.1 Officer-involved Fatalities

Our analysis uses a new county-level panel covering the years 2013-2019,[14] constructed from different data sources. Our main outcome of interest is the difference at the county level between a proxy for the 'true' number of officer-involved fatalities and the officially reported number. Our variable for true homicides is taken from Mapping Police Violence (MPV), a non-governmental, open-source database on lethal police violence. MPV sources its data from other non-governmental websites (Fatal Encounters, The Washington Post), publicly accessible media sources, and publicly available official data sources from local and state agencies.[15]

Counter to this true law enforcement homicide data, we consider officer-involved fatalities in the FBI's Supplementary Crime Report (sourced from the Murder Accountability Project).[16] Alternative official measures of police killings exist, including the National Vital Statistics System (NVSS) data. These include official death records obtained from health agencies. Our decision to use the SHR data rather than the NVSS is primarily grounded in a desire to speak directly to law enforcement agencies' incentives to withhold or manipulate potentially incriminating data. We present the distributions of the MPV and SHR variables and their difference, our main outcome of interest, in Appendix A.

Figure 3 further maps our outcome of interest.[17] It shows the number of years each

[14]We do not include the years past 2019 due to the generalized changes induced by the Covid epidemic.

[15]The MPV data is widely agreed to be the most complete source of true police homicide data (GBD et al., 2021).

[16]In the SHR data, these appear in the homicide category 'Felon killed by police'. Our data include justifiable homicides which refer to the killing of a felon by a peace officer in the line of duty. Both of our measures of officer-involved fatalities include deaths in holding cells but exclude fatalities of inmates in correctional institutions.

[17]We also present separate maps in Figures A1-A2 of Appendix B for MPV and SHR police killings.

county displayed underreporting in the SHR data relative to the MPV data. Many counties show no difference between both police homicide measures in most years. This is mainly because 83.5% of county-year observations display no true police killings at all. Of the 13% of county-years displaying underreporting, 88.3%, show underreporting of 1-2 homicides.[18] However, underreporting is relatively widespread since 43.9% of counties underreport police killings in at least one of the seven years in our data. We provide further descriptive statistics of these outcomes for the US population and different death investigation systems in Table A2 of Appendix F. In general, the underreporting of police killings is lower in coroner counties than in other death investigation systems mainly due to smaller populations.

Figure 4 maps the average report of homicides categorized as 'circumstances undetermined' in the SHR over 2013-2019 by county. As we will show, this homicide category is closely interlinked to underreported police killings. It represents a large share of the total, with 24% of homicides falling under this category among counties that report their SHR data to the FBI.

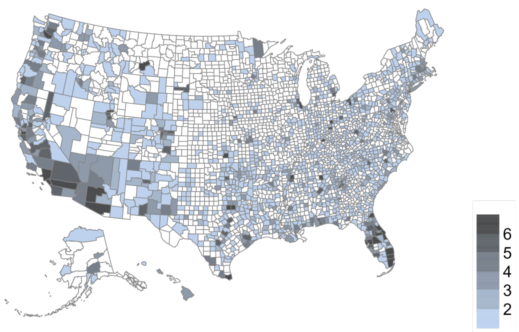


Figure 3: Underreported years

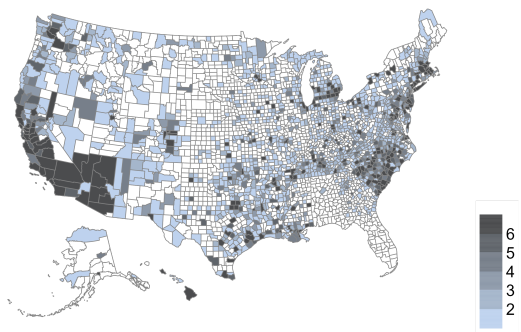


Figure 4: SHR circ-undetermined years

3.2 Law Enforcement Data Reporting

In addition to measures of homicide, we also have yearly information on whether each law enforcement agency shared crime reports to the UCR and NIBRS databases of the FBI.[19] In general, the distribution of law enforcement agencies per county selected in our data is right skewed and varies considerably, with an average of 5.6 law enforcement agencies per county across the US and a maximum of 131 for Cook county, Illinois.[20] The NIBRS,

[18]In a minority of cases, 0.7% of total observations, we find the difference between MPV and SHR data to be negative. In these cases we impute a value of 0 to the difference.

[19]At the time of writing, these data could be downloaded at the following link on the FBI crime explorer website.

[20]To obtain county shares we exclude any agency - such as state troopers, park rangers, university security - which are not defined at the county level. 92% of remaining law enforcement agencies operate only in one county. Of those operating in more than one county, 86% operate in only two counties. For any law enforcement agency spanning multiple counties we divide their participation rate in proportion to the population weight of each county covered.

which has been proposed to replace the UCR program, is designed to include a wider range and more detailed accounts of crimes.[21] Participation rates in the UCR database are high in the years covered by our data, at 86.6%. This UCR information can be seen as a close proxy to the SHR data since only 3.3 of UCR reports do not have associated SHR homicide circumstances. Data sharing to the NIBRS is noticeably lower, at 47.8% and spread unevenly across state lines as a result of some states making the NIBRS mandatory to obtain subsidies in our observation period. For both measures of agency data sharing, we still notice considerable variation within states.[22] In general, medical examiner counties have a higher percentage of agencies sharing to both databases.

3.3 Death Investigation Systems

Our main treatment variable of interest denominating the death investigation system was constructed from multiple sources. We first divide death investigation systems into three mutually exclusive categories, Medical Examiner, Coroner, and Sheriff-Coroner, drawing information from the Center for Disease Control and Hanzlick (2007). For states which allow mixed systems or sheriff-coroners, we searched county by county, consulting county web pages, personal online profiles, Linkedin profiles, and health center records. We also cross-checked the designations of county systems with records from the Census of Medical Examiner and Coroner Offices, 2018[23] and compared our aggregate numbers by state to those presented in Hanzlick (2007). Appendix Figures A5-A8 describe the general distribution of county death investigation systems. Most high population density counties opted to transfer to medical examiner systems during the 1960s-1980s wave of change, but medical examiner counties are also spread throughout more rural areas.

Figure 5 displays the distribution of counties that appointed a sheriff as coroner. The sheriff-coroner system is most prevalent in Nevada, California, and Montana, but is also spread across other areas. Figure 6 shows the distribution of our main treatment variable of interest. It includes sheriff-coroner counties as well as other counties that allow law enforcement to certify a death certificate as defined in Ruiz et al. (2018). We notice that the additional counties which allow law enforcement to certify the cause of death are spread across both medical examiner and coroner counties.

Finally, we draw additional information from census data concerning county populations, race and gender distributions, and county-level GDP from the Bureau of Economic Analysis, as well as presidential election voting share and indicators for the county urbanization level according to the CDC classification. Table A2 in Appendix F offers

[21]Despite being proposed to replace the UCR by Jan. 2021, and some federal programs and grants requiring NIBRS reporting as a condition for funding, the uptake of the NIBRS has been slow. In 2021 only 65% of police departments submitted homicide details in crime reports to the FBI (The Marshall Project, 2022).

[22]We present maps in Figures A3-A4 of Appendix C of the share reporting to the UCR and NIBRS.

[23]This 2018 census was of limited use beyond cross-checking our definitions of death investigation systems as it contains a large share of missing variable values.

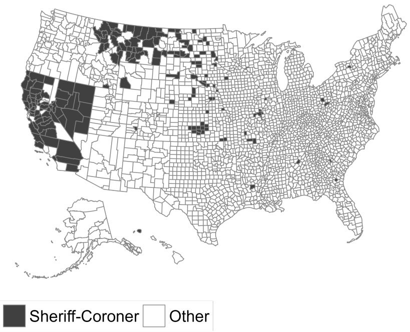


Figure 5: Sheriff-Coroner System

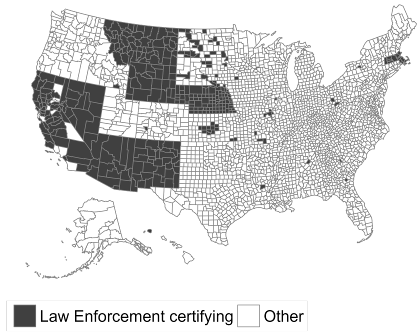


Figure 6: Law Enforcement Certifying

additional descriptive statistics of outcomes and covariates depending on the death investigation system. Of note, we see that Black people are underrepresented in counties where law enforcement is authorized to certify the cause of death relative to the US total, while Hispanic and Native people are slightly overrepresented.

4 Empirical Model

4.1 Causal Analysis Sample

We conduct our study at the level of death investigation systems, which is the county level. The main causal problem we need to address is that counties with different death investigation systems and laws are unequally distributed across the US, and may be different along many dimensions. In addition, police killings can be underreported due to administrative negligence, controversy in the cause of death, or deliberate acts of missclassification and cover up. Since all of these causes are monotonically increasing in the number of homicides, counties with larger populations and higher urban density will mechanically be more likely to display higher levels of underreporting.

Our principal way to address this issue is by taking a fixed effects adjacent treated-control county evaluation approach while additionally controlling for total homicides and other potential confounders. More specifically, in the main analysis, we include only the subsample of counties which allow law enforcement to certify the cause of death, including sheriff-coroner counties, and which neighbor at least one county which does not allow such privileges to law enforcement.[24] Similarly, we include as controls the counties which neighbor at least one county which allows law enforcement to certify the cause of death.[25]

Figure 7 displays the adjacent counties included in our main analysis evaluating the

[24] A similar strategy has previously been used, among others, in Dube et al. (2010) when looking at minimum wage effects.

[25] We provide robustness checks throughout in which we condition the inclusion of adjacent counties on similar categories of urbanization. We also offer robustness checks allowing for second degree adjacency counties to be included.

effect of law enforcement death certification laws. This sample covers 19% of the total US population. Table A2 in Appendix F further presents some descriptive statistics of the relevant treated and control subsamples compared to the full population. In general, we find that outcome and covariate means of law enforcement certifying counties in our analysis sample are representative of law enforcement certifying counties in general. We also find our adjacent control group outcomes and covariates to be representative of US counties as a whole, although the group slightly overrepresents coroner counties and notably underrepresents Black people. We return to both points in our analysis.

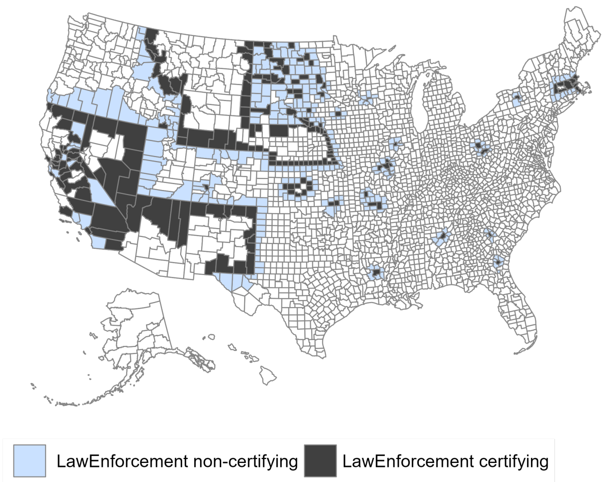


Figure 7: Law Enforcement vs other analysis sample

4.2 Adjacent County Poisson Fixed Effect Model

Our main set of results focuses on the difference between 'true' and reported officer-involved fatalities. Let $\mathrm{Y}_{it} = \mathrm{Y}_{it}^{MPV} - \mathrm{Y}_{it}^{FBI}$ denote the observed underreporting of police killings for county i in year t. We take as our baseline model an adjacent county fixed effect Poisson specification to account for the non-negative and right skewed outcome values. More precisely, denoting by J_i the union of a treated county i and its adjacent control counties, our model is given by,

$$\mathrm{Y}_{it} = \exp\left(\alpha_{J_i} + \lambda_t + \delta\,\mathrm{D}_{it} + \gamma'\,\mathrm{X}_{it}\right)u_{it} \tag{1}$$

In this model, D_{it} is our treatment variable of interest. In our main results, it takes value 1 for counties which allow law enforcement to certify the cause of death, and 0 otherwise. X_{it} includes covariates which control for population, urbanisation, political voting behavior, local economic conditions, and the racial distribution within each county. It also includes the inverse hyperbolic sine transformation of total reported UCR homicides

(H_{it}).[26] α_{J_i} captures additional common unobservables within adjacent county clusters (police culture, violence, etc.). λ_t includes a set of year fixed effects.[27]

δ, the parameter of interest in our model, represents the proportional difference in underreported police killings between counties permitting law enforcement to certify the cause of death and adjacent counties that do not allow law enforcement to do so.

Beyond knowing whether there exists underreporting, we also have an interest in knowing which method is used to hide police killings. One such avenue is simply reclassifying these deaths in other ambiguous crime categories, a possibility we explore using our fixed effects Poisson model. We also offer further evidence of reclassification effects within a new potential outcomes decomposition framework, which we develop within our results section. A second approach to avoid scrutiny is simply to withhold reports of any homicide data. We explore this possibility with our UCR and NIBRS participation data. When estimating the treatment effect on the agency participation data, we specify a linear model with the same adjacent county cluster fixed effects.

4.3 Discussion of Causal Effect and Identification Assumptions

Some assumptions are necessary to allow for a causal interpretation. To discuss these assumptions, we must first clarify the purpose of our evaluation. Police departments in counties which permit (under some circumstances) law enforcement to certify the cause of death have, of course, no obligation of making use of their entitlement. It may even be that if one experimentally randomized a set of non-certifying law enforcement counties to more permissive certification laws, little effect would be noticed in the short run. In the long run, however, differences in legislation may lead to pernicious and collaborative executive, and possibly even judicial, norms. We view the results in this paper as representing these long term differences, enabled by more lenient law enforcement death certification laws. From a policy perspective, we view the measured effect as a close proxy to the change in underreporting of police killings were lenient certification laws to be removed and the sheriff-coroner system to be abolished.

More practically, we evaluate differences in the underreporting of police killings for a subset of neighbouring counties. The fundamental assumption in our evaluation is that these neighbouring counties have equal potential for pernicious cover-up cultures but different death investigation laws and systems which, we argue, are independent of the underreporting outcome. We base this independence assumption on both theoretical and empirical grounds.

As detailed in Section 2.2, major changes in state death investigation systems to

[26] The transformation $log(H_{it} + (H_{it}^2 + 1)^{1/2})$ is defined for $H_{it} = 0$ and is otherwise approximately equal to $log(2H_{it})$. In theory, this variable could be a bad control. We provide substantial evidence in our results that it is not endogenous to our treatment in a discernible way and that it only serves to improve the precision of our estimates.

[27] County fixed effects can not be included since death investigations are fixed at the county level over time.

address the competency of examiners took place between the 1960s-1980s and are largely traceable. The specific history of changes in state death certification laws is less clear. Some of the differences in state laws, such as removing the sheriff-coroner system, came about with the general death investigation system remodeling (Hanzlick, 2007). At the time, aggregated data on police killings was less easily available and the more limited number of media outlets allowed fewer possibilities for public scrutiny. Under these conditions, law enforcement agencies had more leeway in selectively covering up certain killings than they do today. This leeway would presuppose the issue of hiding police killings was not as central in states' decisions to amend death certification laws as was that of providing more competent forensic work.

Other concessions allowing law enforcement to certify the cause of death arose due to state-specific considerations, such as granting more independence in tribal territories or to address delays in rural areas with fewer resources (Ruiz et al., 2018). For these situations, we assume that neighbouring counties divided by a state line are similar along observable and unobservable measures relevant to our study, besides the fact that one county allows law enforcement to certify death certificates due to centrally made state laws, while the neighbouring county does not give such privileges to law enforcement.

Empirically, our analysis also provides a wide set of balance and robustness checks which supports our claim of independence. Speaking indirectly to cultural views on policing, we show in our balance table in Appendix G that demographic and political variables - population, urbanization, racial-ethnic distribution, crime, income - known to be correlated to views on policing culture, including those on voting behavior, are balanced across the treatment and control groups of our analysis in our fixed effect specification.[28] In addition to accounting for possible unobserved differences at the neighbouring county level through adjacency cluster fixed effects, we provide robustness checks including state fixed effects. These produce slightly larger effects than our main ones.

We further produce two insightful analyses on the question of selection looking at the effects for sheriff-coroner counties and other law enforcement certifying counties. The sheriff-coroner analysis mostly exploits within state differences in systems and in some instances differences across state lines, while the non-sheriff coroner analysis exploits a different set of state line divides. Our upcoming analysis shows that the measured effects on the underreporting of police killings are remarkably similar in both analyses despite the fact that only 3% of the sampled counties overlap. These results are in line with our independence assumption since one would expect that, were strong selection at play, isolated sheriff-coroner counties would be more likely to display higher underreporting relative to their neighbouring control counties.

To our knowledge, there were no important changes to legal structures around death investigation systems during our analysis years which were designed with a view to allow

[28]Specifically, we regress each standardized variable on the law enforcement certifying dummy, with (right panel) and without (left panel) adjacency fixed effects.

more permissive police protection in cases of police killings. The Criminal Justice and Forensic Science Reform Act in 2014 was supposed to introduce important changes to the death investigation system, but stalled in the US Senate.

5 Results: Effects of Death Investigation Systems

In this section, we first discuss the main results in our baseline model looking at the effects of permissive death certification laws for law enforcement and sheriff-coroners on the underreporting of police killings. We then move on to discuss how police killings are covered up by being reclassified in other homicide categories. Thereafter, to provide detailed insights into the cover-up behavior, we develop an adapted LATE potential outcomes decomposition method. We then further discuss cover-up methods by looking at whether counties in which law enforcement are permitted to certify the cause of death are more likely to strategically withhold details of their homicide data. Returning to our main results on underreporting, we disaggregate these for different race groups, and go on to specifically discuss the question of the underreporting of illegal immigrant police killings along the US-Mexico border. We further delve into potential roots and solutions to the problem of police killings by focusing on monitoring and threats to police. In a final step, we consider whether our main underreporting effect might be driven by other differences in death investigation systems or death investigators.

5.1 Law Enforcement Certifying Cause of Death

Table 1 presents our main results concerning the effect of allowing law enforcement to certify the cause of death, including sheriff-coroner counties, on police killings. The first column shows the estimate on the underreporting of police killings while only including a covariate for total homicides reported in that county-year in the SHR (scaled in arcsinh). The second column adds our full set of control variables accounting for possible differences in urbanisation levels, political voting behavior, racial distribution, demographics and economic conditions in adjacent counties but does not include any adjacency cluster-specific fixed effects. Both coefficients are positive, showing a $(exp(0.282) - 1) * 100 \approx 32.6\%$ to 40% higher underreported police killings in counties which permit law enforcement to certify the cause of death relative to their controls.

This difference in underreporting increases to 46% in our preferred specification, column 3, which includes adjacent county cluster fixed effects. It represents 0.13 more underreported police killings per year at the mean of 0.29.[29] The effect is of a similar magnitude in column 4 when specifying the adjacent county cluster fixed effect model

[29]Note that the mean in all tables represents the effective estimation mean which drops FE clusters with no variation in outcomes and weights clusters. As a result, these means will differ from those presented in summary statistics presented in Appendix G.

with a linear specification. We find counties that allow law enforcement to certify death have 0.09 more underreported police killings than their adjacent controls.

Table 1: Baseline effects of authorising law enforcement to certify cause of death

Dep. Var.:	Diff. L.E. homicides					
	(1)	(2)	(3)	(4)	(5)	(6)
LE certify	0.326** 0.282 (0.119) [0.018]	0.402*** 0.338 (0.101) [0.001]	0.459*** 0.378 (0.122) [0.002]	0.091*** (0.033) [0.006]		
Sher-Coroner					0.689*** 0.524 (0.193) [0.007]	
LE certify No Sher.-Coroner						0.732** 0.549 (0.220) [0.012]
All covariates	No	Yes	Yes	Yes	Yes	Yes
FE	No	No	Yes	Yes	Yes	Yes
Spec.	Pois.	Pois.	Pois.	Lin.	Pois.	Pois.
N_{tot}	3,464	3,464	3464	3464	2128	1238
N_{eff}	3,464	3,464	2736	3464	1666	943
N_{treat}	1280	1280	930	1280	497	335
$N_{counties}$	495	495	391	495	238	135
$\mu_{outc.}$	0.228	0.228	0.289	0.228	0.315	0.262

Note: Table displays transformed coefficient with *p<0.1; **p<0.05; ***p<0.01, followed by estimated coefficient, its standard error, and its p-value. Poisson estimation with year and adjacency county cluster fixed effects, and standard errors clustered at the county level. Set of controls described in Appendix G.

In Table A4 of Appendix H we provide several additional robustness checks to address potential confounding: (1) we estimate a specification which includes adjacent county-cluster fixed effects and state fixed effects, dropping any state with no variation in our law enforcement certification variable of interest, (2) we estimate a specification which includes a covariate indicating whether the death investigator was a Physician,[30] (3) we estimate a specification which drops the potential bad control *total homicide* variable, (4) we add to our main specification a gun law covariate capturing the stringency of gun and ammunition laws by state,[31] (5) we provide a specification excluding off-duty officer

[30]This is potentially a bad control, as it is obviously correlated with not being a sheriff-coroner. Our later evaluations of other characteristics of death investigation systems on different adjacent samples, speaking more directly to these potential confounding effects, show, somewhat surprisingly, no effects of medical qualifications.

[31]The gun law variable is a yearly measure ranging from 0-11 based on Giffords law center scorecard, where 11 corresponds to a score of A and 0 to a score of F. The yearly state score weighs the many laws pertaining to ammunition and gun distribution, possession, and right-to-carry within a state. We again do not include this covariate in most of our specifications since it could, potentially, be a bad control.

police killings from the MPV police killings, of which there are only 33 out of the 3,146, since these may not always be officially reported in the SHR as police killings, (6) we account for different possible inclusion criteria by looking at the underreporting outcome of *MPV police killings by firearm - SHR police killings*,[32] (7) we consider a model which offsets on the total homicide variable instead of controlling for it, (8) we consider a model which offsets on true MPV police killings, (9) we estimate our preferred specification but only matching treated and control counties with the same urbanization level, (10) we produce estimates when allowing for second degree adjacency counties to be included, (11) we produce estimates when weighting each county-year observation in the sample by the propensity of entering the sample, which gives an approximate extrapolation to the effect on the entirety of the US. The estimates from these different specifications range from 40%-80%. In consequence, we consider the stated result of 46% from our preferred specification to be a conservative estimate.

In addition, the specification of column 11 allows us, using a back-of-the-envelope calculation, to extrapolate the approximate long-term reduction in underreported police killings were the law enforcement certification laws abolished throughout the US. Abolishing these laws would produce a reduction of 316 underreported killings every 10 years. It should be noted that this number is but a minor representation of the potential consequences of abolishing death certification laws inhibiting police accountability. The social consequences are likely more far-reaching.

The final two columns of Table 1 separate effects for sheriff-coroner counties and counties in which law enforcement can certify the cause of death but which are not sheriff-coroner counties. For the sheriff-coroner analysis of column 5 and the analysis of other law enforcement certifying counties in column 6, the adjacent control counties exclude any law enforcement certifying county. Despite an overlap of only 3% of counties, the effects measured in both samples are of the same order of magnitude, 68.4% and 73.2%. This result shows that the main effects measured in column 3 are not solely driven by sheriff-coroner counties. As discussed in Section 4.3, it also lends credibility to our independence assumption. As in the additional robustness checks in Appendix H, these separate effects are also larger than in their combined analysis of column 3.

5.2 Strategic Re-classification

Table 2 delves further into the mechanisms that are used to cover up police killings. Columns 1 and 2 present the separate parts of our underreporting outcome, namely the

[32] 'Justifiable homicides' should in principle only include deaths resulting from situations in which an officer intended to use lethal force (Zimring, 2017). In practice, the line of what falls under this definition is murky. Death by 'Taser', 'Physical restraint', 'Beaten', 'Asphyxiated', among others, may or may not be included. Ultimately, this qualification constitutes one of the margins which offers room to avoid official reporting, and is integral to the analysis. Even so, we still consider the specification in which we only include MPV deaths by firearm, which is a strong indication of potential lethal use of force, as the 'true' police killing variable compared to SHR reported deaths.

reported SHR and true MPV officer-involved fatalities respectively. We find evidence for both lower reported SHR police killings and higher MPV police killings in counties permitting law enforcement to certify the cause of death compared to their controls. Although neither is significantly different from 0, the difference in MPV police killings may indicate that permissive laws aiding the cover-up of police killings contribute to a culture of more police killings.

Figure 8 further touches on the question of violent policing culture looking at sub-circumstances of SHR police killings (column 1 of Table 2). Sub-circumstances are listed for 89% of SHR police killings. Estimating linear probability models, the results indicate that police killings in law enforcement certifying counties occur less often in response to attacks by victims and more often when a victim was fleeing or under undisclosed reasons. If taken at face value, these results are insightful. Since the Supreme Court decision in Tennessee v. Garner (1985), deadly force should only be used on a fleeing suspect if that suspect committed egregious crimes or poses an immediate threat to others.[33] Our results indicate that officers in law enforcement certifying counties apply a less strict threshold when choosing to use lethal violence against fleeing offenders. The result showing fewer justified reasons for lethal violence ('Not enough info') may also suggest officers assume they are likely to be questioned when filing police homicide reports.

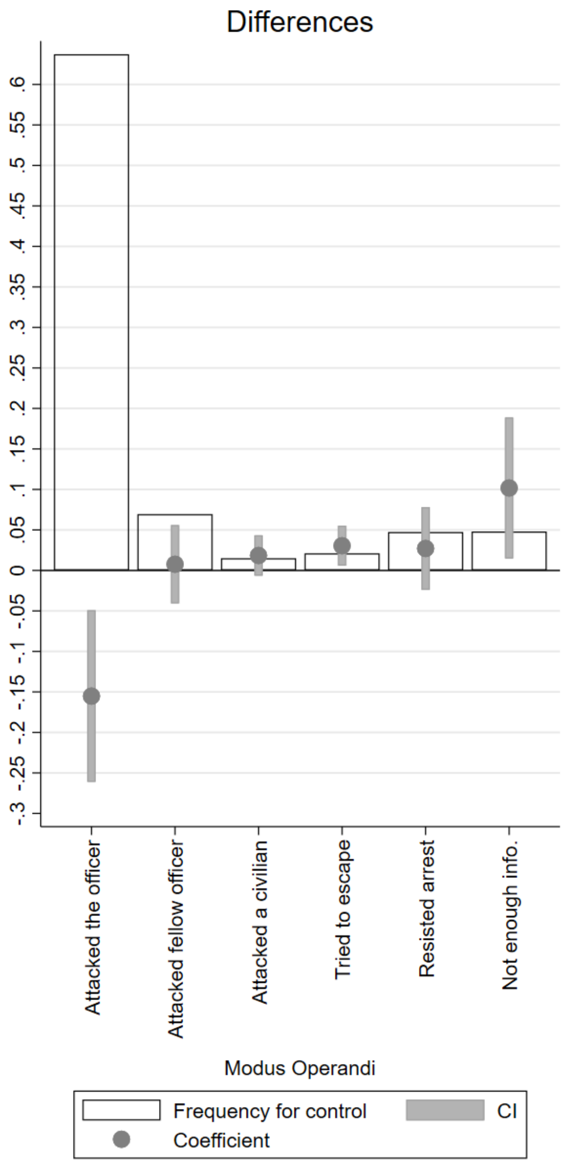


Figure 8: Effects on SHR sub-circumstances of death

Returning to the results of Table 2 , the MPV police killing data also indicates whether the deceased person had mental health issues. This information is gathered from news and official reports. According

[33]There is some correlational evidence that this led to a slight reduction in reported police killings (Tennenbaum, 1994), though in light of our results, without crowed-sourced 'true' numbers of police killings that go so far back in time, it is unclear whether those changes were due to increased underreporting.

to the MPV data, approximately 19% of victims of police killings in the US had mental health problems at the time of the homicide or had a history of mental health issues. Column 3 of Table 2 evaluates whether permissive death certification laws for police influence the reporting of mental health issues. We find some indication, marginal at customary significance levels, that law enforcement certifying counties have higher reports of police killings where the victim displayed or had a history of mental health issues.

Because the source of the mental health information can be official or from news reports, there are different ways to interpret these results, were they to be true. If the effect is largely driven by official reports, then law enforcement agencies in counties with permissive death certification laws may be manipulating reports to justify police killings as necessary due to the unpredictable mental health state of the victim. If the main source of the effect is driven by independent news reports, which would better reflect the true state of victim's mental health at the time of killing, then the measured effect may imply that agents in law enforcement certifying counties have a lower inclination to explore less lethal moderation methods when it comes to individuals showing signs of mental illness. These results may also show that the media highlight exonerating factors in police killings, a finding that would parallel text analysis results in Moreno-Medina et al. (2022), also using the MPV dataset, with regard to the media's obfuscatory language in cases of police killings.

In general, it is of course difficult to hide a homicide entirely, even if, in principle, a law enforcement officer might tamper with evidence and report a suicide[34] or a natural death.[35] However, only a minority of killings are not by police gunshot ($\sim 6\%$) and therefore could potentially be reclassified as natural deaths or suicides.

An easier way to hide these homicides, especially when they are by gunshot, is to reclassify them in an alternative homicide category. Column 4 of Table 2 investigates whether counties that allow law enforcement agents to certify the cause of death show a higher reporting of homicides categorised as 'circumstances undetermined', the catch-all residual circumstance of homicide category in the SHR. We find that law enforcement certifying counties have 35.1% more 'circumstances undetermined' homicides than their adjacent controls, supporting a theory of strategic reclassification. Columns 5 and 6 separate these effects for sheriff-coroner and other law enforcement certifying counties. Our results indicate that reclassification into 'circumstances undetermined' is fully driven by the 60% of law enforcement certifying counties that have a sheriff-coroner (column 5), with no significant effect for the remaining 40% of treated counties (column 6).

Column 7 additionally shows that total UCR homicides in law enforcement certifying counties are no different from those in control counties. This result indicates that the higher reported 'circumstances undetermined' homicides in law enforcement certifying

[34] Unfortunately, for privacy reasons, suicide data at the county-year level are difficult to access.

[35] In the case of George Floyd, the defense attempted to call upon 'excited delirium', a - non-testable, unproven in humans - cause of death in which an excessive amount of stress on certain animals causes sudden death.

Table 2: Officer-involved fatalities and Re-classification effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LE certify	-0.084	0.139	0.409*	0.351***			0.021	0.050	0.026***
	-0.088	0.130	0.343	0.301			0.021	0.049	0.026
	(0.103)	(0.093)	(0.184)	(0.076)			(0.085)	(0.040)	(0.011)
	[0.396]	[0.161]	[0.062]	[0.000]			[0.806]	[0.222]	[0.022]
Sher-Coroner					0.372***				
					0.316				
					(0.083)				
					[0.000]				
LE certify No Sher.-Cor.						-0.236			
						-0.269			
						(0.252)			
						[0.287]			
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spec.	Pois.	Pois.	Pois.	Pois.	Pois.	Pois.	Pois.	Pois.	Pois.
N_{tot}	3464	3464	3464	3464	2128	1238.00	3,464	3,464	355
N_{eff}	2736	2736	1749	2820	1904	803.00	3,366	3,366	350
N_{treat}	930	930	580	937	553	293.00	1196	1196	248
$N_{counties}$	391	391	250	403	272	115.00	481	481	105
$\mu_{outc.}$	0.326	0.607	0.405	2.122	2.411	1.720	5.427	5.158	40.783

Note: Table displays transformed coefficient with $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$, followed by estimated coefficient, its standard error, and its p-value. Dependent variable: (1) SHR police killings, (2) MPV police killings, (3) MPV police killings with reported mental health issues, (4) SHR: circumstances undetermined, (5) SHR: circumstances undetermined, sheriff-coroner adjacency sample, (6) SHR: circumstances undetermined, non-sheriff-coroner LE certify adjacency sample, (7) SHR: Total homicides, (8) SHR: Total homicides - police killings, (9) SHR: Total homicides - police killings, conditional on positive MPV police killings. Poisson estimation with county adjacency year and adjacency county cluster fixed effects, and standard errors clustered at the county level. Set of controls described in Appendix G.

counties is not simply due to higher overall homicides in those counties.[36]

Columns 8 and 9 further consider whether law enforcement certifying counties appear to have higher homicides in the remainder of their total UCR homicides after excluding the SHR police killings from column 1. Column 8 presents differences in remainder homicide reporting for all adjacent counties, while column 9 shows these differences conditioning on counties which actually show positive MPV true killings. We find higher non-police killing homicides reported in both columns, but these are only significant in the specification conditioning on counties actually observed to have true police killings. Taken together, the effects in Table 2 offer substantial evidence that there are higher homicides specifically in the 'circumstances undetermined' category for counties which allow law enforcement to certify the cause of death.

5.3 Decomposing Total Effects for Compliers, Never-takers and Always-takers

The evidence for re-classification as a cover-up mechanism seems convincing given the average effects on underreported police killings combined with those on 'circumstances undetermined' homicides. But several questions remain given these observed effects. First, we would like to know whether the effects are driven by moderate differences among many counties, or by large differences among a few of them. In addition, we are interested in whether the average effects are driven by counties in which law enforcement agencies would only underreport police killings when facing lenient certification laws. They may also be due to counties in which a large share of law enforcement agencies would always underreport police killings, but would use different cover-up strategies depending on the certification laws. To address these questions, among others, concerning which population groups contribute to our average effects, we propose a modified LATE potential outcomes framework inspired by Imbens and Angrist (1994) which also draws parallels with the discussion of supercompliers in Comey et al. (2022).

5.3.1 Identification of Types

For simplicity of presentation, we take a one-period model and ignore the time t subscript. We also leave the conditioning on exogenous covariates X_{it} implicit since identification is nonparametric. Take D_i as the treatment status for agent i, $i = 1, \dots, N$. In our case, $\mathrm{D}_i = 1$ for counties in our adjacent sample in which law enforcement can certify the cause of death, and $\mathrm{D}_i = 0$ otherwise. In addition, we define the potential outcomes for two variables Y_i^d and W_i^d, $d = 0, 1$. Here, we consider only the case in which Y_i^d and W_i^d are both binary.[37] In our study, $\mathrm{Y}_i^d = \mathbb{1}[\mathrm{Y}_{i,mpv}^d - \mathrm{Y}_{i,shr}^d > 0]$ takes value 1 if county

[36] This result also offers empirical evidence that total UCR reported homicides are not a bad control in our main specification.

[37] Without additional assumptions than those presented in this paper, it is unclear whether a causal decomposition can be identified with non-binary outcomes.

i underreports police killings when exposed to $\mathrm{D}_i = d$, and 0 otherwise. Similarly, take $\underline{\mathrm{W}}_i^d$ as the ratio of reported 'circumstances undetermined' homicides to total reported UCR homicides when exposed to $\mathrm{D}_i = d$. Then $\mathrm{W}_i^d = \mathbb{1}[\underline{\mathrm{W}}_i^d > median(\underline{\mathrm{W}}_i^0)]$ takes value 1 if county i reports $\underline{\mathrm{W}}_i^d$ when exposed to $\mathrm{D}_i = d$ above the sample median when exposed to $\mathrm{D}_i = 0$, and 0 if $\underline{\mathrm{W}}_i^d$ is equal or lower than the median ratio under $\mathrm{D}_i = 0$. We refer henceforth to W_i^d as the 'excess undetermined homicides' variable. Figures 9 and 10 display the number of years in which each county underreported police killings and reported excess undetermined homicides. The overlap in both outcome variables is noticeable.

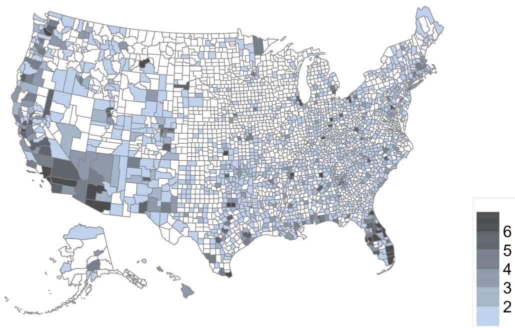


Figure 9: Underreported police killings

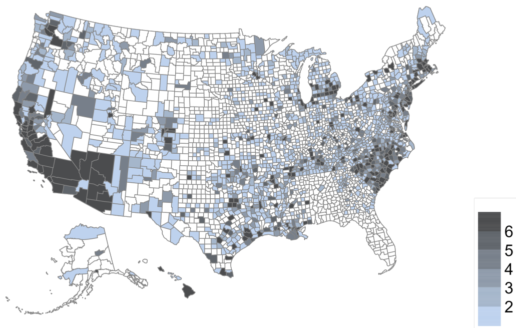


Figure 10: Excess undetermined homicides

With this notation in hand, we can define probabilities representing the possible combinations of potential outcomes. These are defined, omitting for convenience the i subscript, by the joint probability $\Pr(\mathrm{W}^1 = w, \mathrm{W}^0 = w^*, \mathrm{Y}^1 = y, \mathrm{Y}^0 = y^*)$ for each of w, w^*, y, y^* equal to 0 or 1. For example, we are particularly interested in the probability that a county synchronously displays underreported police killings and excess undetermined homicides only when subject to laws permitting law enforcement to certify the cause of death. The probability that a county belongs to this group is $\Pr(\mathrm{W}^1 = 1, \mathrm{W}^0 = 0, \mathrm{Y}^1 = 1, \mathrm{Y}^0 = 0)$. Borrowing from the LATE nomenclature, this probability can be understood as that of the W-complier and Y-complier county type, or principal strata, which we write more succinctly as $\Pr(c^W c^Y)$. We can similarly define always-taker, never-taker, and defier county types for both outcomes of interest, resulting in a total of 16 probabilities representing each combination of county types. Each of these types is described in Table A5 of the Appendix I.

Our goal is to identify a maximal amount of these 16 probabilities given a set of plausible behavioral assumptions about law enforcement's reaction to changes in death certification laws. Our first two assumptions draw parallels to those of Imbens and Angrist (1994) and Comey et al. (2022), with the difference being that we are not only considering classical instrument-treatment-outcome settings.

Assumption A.I: Unconfoundedness:

$$(\mathrm{Y}_i^1, \mathrm{Y}_i^0, \mathrm{W}_i^1, \mathrm{W}_i^0) \perp\!\!\!\perp \mathrm{D}_i$$

Assumption A.II: Outcome Monotonicity: *For all i*

$$(i) \quad \Pr(\mathrm{Y}_i^1 \geq \mathrm{Y}_i^0) = 1$$
$$(ii) \quad \Pr(\mathrm{W}_i^1 \geq \mathrm{W}_i^0) = 1$$

Assumptions A.I is a classical unconfoundedness assumption. It requires that in our adjacent analysis sample all potential outcomes of interest are random with respect to whether law enforcement is permitted to certify the cause of death. A.II draws parallels with Comey et al. (2022) by proposing two monotonicity assumptions on the outcomes of interest. The first excludes the existence of counties in which law enforcement agencies would underreport police killings when not able to certify the cause of death but would not underreport police killings when facing more permissive laws for certifying causes of death. The second part, A.IIii, adds the same monotonicity assumption to the excess undetermined homicide outcome.[38]

Assumption A.II fixes any probability with a defier type to equal zero, reducing 16 unknown probabilities to 9. Under assumption A.I and A.II, we can already identify the shares of each county type depending on the outcome,

$$\begin{aligned}
&\Pr(\mathrm{W}^1 = 1, \mathrm{W}^0 = 1) \equiv \Pr(a^W) = \Pr(\mathrm{W} = 1|\,\mathrm{D} = 0)\\
&\Pr(\mathrm{W}^1 = 1, \mathrm{W}^0 = 0) \equiv \Pr(c^W) = \Pr(\mathrm{W} = 1|\,\mathrm{D} = 1) - \Pr(\mathrm{W} = 1|\,\mathrm{D} = 0)\\
&\Pr(\mathrm{W}^1 = 0, \mathrm{W}^0 = 0) \equiv \Pr(n^W) = 1 - \Pr(\mathrm{W} = 1|\,\mathrm{D} = 1) = \Pr(\mathrm{W} = 0|\,\mathrm{D} = 1)\\
&\Pr(\mathrm{Y}^1 = 1, \mathrm{Y}^0 = 1) \equiv \Pr(a^Y) = \Pr(\mathrm{Y} = 1|\,\mathrm{D} = 0)\\
&\Pr(\mathrm{Y}^1 = 1, \mathrm{Y}^0 = 0) \equiv \Pr(c^Y) = \Pr(\mathrm{Y} = 1|\,\mathrm{D} = 1) - \Pr(\mathrm{Y} = 1|\,\mathrm{D} = 0)\\
&\Pr(\mathrm{Y}^1 = 0, \mathrm{Y}^0 = 0) \equiv \Pr(n^Y) = 1 - \Pr(\mathrm{Y} = 1|\,\mathrm{D} = 1) = \Pr(\mathrm{Y} = 0|\,\mathrm{D} = 1)
\end{aligned} \tag{2}$$

In addition, we can relate the nine remaining county type probabilities to observed outcomes as follows:

$$\begin{aligned}
&\Pr(W = 1, Y = 1|D = 0) = \Pr(a^W a^Y)\\
&\Pr(W = 1, Y = 1|D = 1) = \Pr(a^W a^Y) + \Pr(a^W c^Y) + \Pr(c^W a^Y) + \Pr(c^W c^Y)\\
&\Pr(W = 1, Y = 0|D = 0) = \Pr(a^W c^Y) + \Pr(a^W n^Y)\\
&\Pr(W = 1, Y = 0|D = 1) = \Pr(a^W n^Y) + \Pr(c^W n^Y)\\
&\Pr(W = 0, Y = 1|D = 0) = \Pr(n^W a^Y) + \Pr(c^W a^Y)\\
&\Pr(W = 0, Y = 1|D = 1) = \Pr(n^W a^Y) + \Pr(n^W c^Y)\\
&\Pr(W = 0, Y = 0|D = 1) = \Pr(n^W n^Y)
\end{aligned}$$

The above effectively presents us with a system of five equations and seven unknowns. We therefore require an additional assumption to identify the remaining type probabilities. One such assumption, as presented in the identification framework of Comey et al.

[38]Strictly speaking, because 90% of counties comprise of multiple law enforcement agencies, the assumptions must only hold on aggregate at the county level.

(2022), would be an exclusion restriction on the influence of death certification laws. Unfortunately, their exclusion restriction assumption is not plausible in our setting. It would impose that declaring excess undetermined homicides depends only on whether there were underreported police killings and not on whether the county was facing lenient death certification laws. This is unlikely to hold since police departments would presumably use different cover-up methods depending on the leniency of death certification laws.

Instead, we elicit the following two assumptions consistent with incentives in our setting and which we further justify in our empirical results,

Assumption A.III: Cross-Monotonicity Restriction: *For all i*

$$(i) \quad \Pr(\mathrm{W}_i^1 = 1, \mathrm{W}_i^0 = 0, \mathrm{Y}_i^1 = 0, \mathrm{Y}_i^0 = 0) \equiv \Pr(c^W n^Y) = 0$$
$$(ii) \quad \Pr(\mathrm{W}_i^1 = 0, \mathrm{W}_i^0 = 0, \mathrm{Y}_i^1 = 1, \mathrm{Y}_i^0 = 0) \equiv \Pr(n^W c^Y) = 0$$

Assumption A.IIIi and A.IIIii are tailored to our setting but other combinations of 0-probability types can also be used to achieve identification.[39] Assumption A.IIi states that there are no counties in which, on aggregate, law enforcement agencies would not respond to permissive certification laws by increasing the underreporting of police killings, but would respond to these laws by increasing reported excess undetermined homicides. This seems plausible as long as the main reason for increasing declared excess undetermined homicides in response to more permissive certification laws is to hide certain police killings.

Assumption A.IIIii states that there are no counties in which, on aggregate, law enforcement agencies would respond to permissive certification laws by increasing the underreporting of police killings, but would not respond to these laws by increasing reported excess undetermined homicides. This assumption effectively claims that the prevalent cover-up strategy for underreporting law enforcement agencies in complier counties is to reclassify police killings as 'circumstances undetermined' homicides.[40] If this is the case, then underreported police killings specifically induced by the more permissive laws will always show up on aggregate at the county level as excess undetermined homicides.

Assumptions A.I-A.III allow us to identify all remaining county-type probabilities. We present the proof in Appendix I. With $\Pr(a^W a^Y)$ and $\Pr(n^W n^Y)$ already defined, the

[39] Five probability types can be identified under any combination setting one of $\Pr(c^W n^Y)$, $\Pr(a^W n^Y)$ or $\Pr(a^W c^Y)$ to 0 with setting any one of $\Pr(n^W c^Y)$, $\Pr(n^W a^Y)$ or $\Pr(c^W a^Y)$ to 0.

[40] In addition, as elaborated on in Section 5.4, we must also assume that not all underreporting complier law enforcement agencies which reclassify police killings opt to withhold sharing their SHR homicide data to the FBI.

remaining type probabilities are given by,

$$
\begin{aligned}
\Pr(c^W c^Y) &= \Pr(W=1, Y=1|D=1) - \Pr(W=1, Y=1|D=0) \\
&\quad + \Pr(W=1, Y=0|D=1) - \Pr(W=1, Y=0|D=0) \\
&\quad + \Pr(W=0, Y=1|D=1) - \Pr(W=0, Y=1|D=0) \\
&= \Pr(W=0, Y=0|D=0) - \Pr(W=0, Y=0|D=1) && \textit{upper} \\
\Pr(a^W c^Y) &= \Pr(W=1, Y=0|D=0) - \Pr(W=1, Y=0|D=1) && \textit{lower} \\
\Pr(a^W n^Y) &= \Pr(W=1, Y=0|D=1) && \textit{upper} \\
\Pr(c^W a^Y) &= \Pr(W=0, Y=1|D=0) - \Pr(W=0, Y=1|D=1) && \textit{lower} \\
\Pr(n^W a^Y) &= 1 - \Pr(W=1, Y=1|D=1) && \textit{upper}
\end{aligned}
\tag{3}
$$

In addition, as noted, if either of Assumptions A.IIIi and A.IIIii does not hold, we can still imply potentially informative upper and lower bounds on the above type probabilities. From the joint probabilities in equations 3 and the share probabilities in equations 2, we can also apply Bayes' rule to obtain all conditional probabilities of interest such as $\Pr(\mathrm{W}^1 = 1, \mathrm{W}^0 = 0|\, \mathrm{Y}^1 = 1\, \mathrm{Y}^0 = 0) = \Pr(c^W c^Y)/\Pr(c^Y)$.

5.3.2 Type Probability Estimates

Our main interest in decomposing the total effect into type probabilities is knowing which population groups contribute to our average effects. Table 3 therefore presents results for all probabilities that include complier counties, those induced to change outcome(s) when facing more lenient certification laws. Each column estimates a probability change using a linear model with adjacency cluster fixed effects.[41]

Column 1 presents the average effect of lenient law enforcement certification laws on our constructed underreporting binary outcome, $\Pr(Y=1|D=1) - \Pr(Y=1|D=0)$. The estimates indicate 4.9 percentage point higher underreporting of police killings in these counties.

We can further decompose this total effect into two joint probability effects. The first joint probability effect of $\Pr(W=1, Y=1|D=1) - \Pr(W=1, Y=1|D=0) = \Pr(a^W c^Y) + \Pr(c^W a^Y) + \Pr(c^W c^Y)$ is presented in column 2 of Table 3. It shows the effect of permissive law enforcement death certification laws on the probability that a sufficiently large share[42] of law enforcement agencies are jointly induced to underreport police killings and induced to report excess circumstances undetermined homicides. Our results indicate 4.3 percentage point more underreporting in law enforcement certifying counties.

The results from columns 3-4 further qualify the results of column 2. Under assumptions I-III, column 3 estimates $-\Pr(c^W a^Y) = \Pr(W=0, Y=1|D=1) - \Pr(W=0, Y=$

[41]We further discuss data related issues and present additional results in appendix I when generating probabilities from logit specifications.

[42]By this we mean that enough law enforcement agencies within the county change their reporting behavior such that our outcomes, Y and/or W, depending on the analysis, shift from 0 to 1.

$1|D = 0)$, which is the second joint probability effect contributing to the effect of column 1. Column 4 estimates $-\Pr(a^W c^Y) = \Pr(W = 1, Y = 0|D = 1) - \Pr(W = 1, Y = 0|D = 0)$. According to the results, the effects for $\Pr(c^W a^Y)$ and $\Pr(a^W c^Y)$ are not significantly different from zero, and relatively tightly bounded. This implies that the total underreporting effect in column 1 is largely (88%) driven by counties with higher underreporting of police killings and, suspiciously, simultaneous increases in reported circumstances undetermined homicides, the $\Pr(c^W c^Y)$ group.

Beyond their intrinsic value, these synchronous effects reinforce our initial results on law enforcement certification laws. Any alternative explanation for our measured increases in the underreporting of police killings due to law enforcement certification laws would need to simultaneously explain the increase in 'circumstance undetermined' homicides. One such explanation might be to postulate that higher media scrutiny in sheriff-coroner counties results in higher reported MPV true police killings, and as a result more measured underreporting in sheriff-coroner counties. This explanation, however, cannot explain the simultaneous increase in excess undetermined circumstances homicides.

The results also partially test Assumption III. Note that significant positive treatment effects in columns 3 and 4 results would imply that $Pr(c^W n^Y) > 0$ or $Pr(n^W c^Y) > 0$. Given that we find a zero treatment effect, it is unlikely for counties to be simultaneously compliers and never-takers, unless it so happens that $Pr(c^W n^Y) = Pr(a^W c^Y) \neq 0$ and $Pr(n^W c^Y) = Pr(c^W a^Y) \neq 0$.

Table 3: Underreporting and Undetermined Causes of Death

Dep. Var.:	Underreport	Underreport & Undeterm.	Underreport & Determ.	Report & Undeterm.
	$Y = 1$	$Y = 1$ & $W = 1$	$Y = 1$ & $W = 0$	$Y = 0$ & $W = 1$
	(1)	(2)	(3)	(4)
LE certify	0.049***	0.043**	0.006	-0.007
	(0.012)	(0.010)	(0.021)	(0.008)
	[0.000]	[0.000]	[0.781]	[0.404]
Covariates	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes
Spec.	Lin.	Lin.	Lin.	Lin.
N_{tot}	3464	3464	3464	3464
N_{eff}	3464	3464	3464	3464
N_{treat}	1280	1280	1280	1280
$N_{counties}$	495	495	495	495
$\mu_{outc.}$	0.120	0.079	0.399	0.042

Note: Table displays coefficient with *p<0.1; **p<0.05; ***p<0.01, followed by standard error, and p-value. Panel linear model with year and adjacency county cluster fixed effects. Set of controls described in Appendix G.

Given that the results do nt reject Assumption A.III, we proceed by estimating the other-type probabilities and presenting them in Appendix I. These remaining type probabilities are estimated using different specifications. Some results are worth highlighting. We find that around 85% of counties never underreport police killings, and about 10% of counties always underreport them. We further show that, conditional on underreporting police killings, a county will almost surely report excess undetermined homicides. Similarly, conditional on reporting excess undetermined homicides, a county will almost surely underreport police killings.

It should be noted that, if assumption A.III were incorrect, allowing $\Pr(c^W n^Y) > 0$ and $\Pr(n^W c^Y) > 0$, then $\Pr(c^W c^Y)$ would be smaller than our estimated effect while $\Pr(a^W c^Y)$ and $\Pr(c^W a^Y)$ would be larger than our estimated effects. We discuss these bounds in relation to our results in more detail in Appendix I.

5.4 Strategic Withholding of Data

Besides reclassifying police killings into other homicide categories, an additional or alternative strategy would be for law enforcement agencies to avoid sharing homicide information altogether with the FBI. Table 4 explores this possibility using the data on agency reporting. Column 1 considers whether county-level participation rates of law enforcement agencies in the FBI's UCR program, a prerequisite to the SHR participation analyzed in previous tables, differ between law enforcement certifying counties and their controls.

A priori, the expected effect on UCR participation is unclear. When facing restrictive death certification laws, some law enforcement agents and agencies may be less able to reclassify police killings as 'circumstances undetermined' homicides. As a result, they may decide to avoid scrutiny by not sharing UCR data entirely, which would predict a negative effect of permissive death certification laws on UCR participation. In contrast, it may be that some law enforcement agencies which reclassify police killings as documented in previous tables will additionally cover their tracks by not sharing UCR data. The results in column 1 suggest that this second theory prevails on aggregate. We find that law enforcement certifying counties are 3.8 percentage points less likely to share their crime data with the UCR.

Column 2 looks further into differences in the sharing of detailed homicide data by considering submissions of reports to the NIBRS. The results indicate that law enforcement agencies in counties with more lenient death certification laws are 18.2 percentage points less likely to submit a report of the detailed circumstances of death.[43]

Columns 3-6 examine these differences in UCR and NIBRS participation rates for sheriff-coroner and other law enforcement certifying counties separately. Column 7 adds to these results by including state fixed effects in the estimation of NIBRS reporting for

[43]When considering whether law enforcement agencies in counties report the UCR or NIBRS, our results are practically identical to the UCR results of column 1.

non-sheriff law enforcement certifying counties. Our results indicate that differences in participation to both the UCR and NIBRS are mainly driven by counties which permit law enforcement to certify the cause of death but are not sheriff-coroner counties. Considering these results in hand with those of Tables 2 and 3, we surmise that the main cover-up method for police departments is to reclassify police killings into the 'circumstances undetermined' homicide category. In addition, law enforcement agencies in sheriff-coroner counties do so without additionally resorting to hiding their homicide data from the public. This less cautious behavior is consistent with the theory that sheriff-coroners, united with police departments behind the 'blue wall of silence', are more likely to stand behind the reclassified cause of death in the event of an external inquiry.

Table 4: Law enforcement agency data sharing to UCR and NIBRS

Dep. Var.:	UCR	NIBRS	UCR	NIBRS	UCR	NIBRS	NIBRS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LE certify	-0.038**	-0.182***					
	(0.018)	(0.028)					
	[0.039]	[0.000]					
Sher-Coroner			-0.028	-0.039			
			(0.025)	(0.025)			
			[0.275]	[0.115]			
LE certify					-0.067**	-0.364***	-0.396***
No Sher.-Cor.					(0.029)	(0.051)	(0.150)
					[0.024]	[0.000]	[0.008]
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes + St.
Spec.	Lin.	Lin.	Lin.	Lin.	Lin.	Lin.	Lin.
N_{tot}	3,464	3,464	2128	2128	1238	1238	1238
N_{eff}	3,464	3,464	2128	2128	1238	1238	1238
N_{treat}	1280	1280	644	644	510	510	510
$N_{counties}$	495	495	304	304	177	177	177
$\mu_{outc.}$	0.874	0.511	0.882	0.495	0.858	0.540	0.540

Note: Table displays coefficient with *p<0.1; **p<0.05; ***p<0.01, followed by standard error, and p-value. Panel linear model with year and adjacency county cluster fixed effects. Column 7 includes state fixed effects as well. Standard errors clustered at the county level. Set of controls described in Appendix G.

5.5 Death Investigation Systems and Race

5.5.1 Underreporting Effects by Race

Table 5 considers whether the effects uncovered for counties allowing law enforcement to certify the cause of death in Table 1 are different by race and ethnicity. The race-ethnicity groupings are chosen by necessity as they are the only ones which allow a direct comparison of the MPV and SHR police killings data. In the top frame we present results from our adjacent county fixed effect Poisson specification. Fixed effects Poisson has a drawback when it comes to analysing effects for different race groups. In our data, officer involved homicides of any specific race, besides White people, are relatively low in absolute terms. This means that many counties will be excluded in the Poisson estimation which drops any fixed effect cluster of adjacent counties with 0 outcomes in all years. To account for this low observation problem, we also present results from the Poisson specification without fixed effects in our lower panel. The latter will not account for possible unobserved heterogeneity at the county adjacency cluster.

Table 5: Race Group Effects

Dep. Var.:	Diff. L.E. homicides			
Victims:	White	Non-White	Black	Hispanic
	(1)	(2)	(3)	(4)
Specif.:	FE Poisson model			
LE certify	0.537***	0.079	-0.034	0.283
	0.43	0.076	-0.035	0.249
	(0.141)	(0.177)	(0.276)	(0.251)
	[0.002]	[0.668]	[0.898]	[0.320]
N_{tot}	3,464	3,464	3464	3464
N_{eff}	2,435	1,820	826	1176
N_{treat}	811	651	315	476
$N_{counties}$	348	260	118	168
$\mu_{outc.}$	0.132	0.244	0.127	0.244
Specif.:	Poisson model			
LE certify	0.401**	0.300*	0.130	0.645**
	0.337	0.262	0.122	0.498
	(0.143)	(0.151)	(0.251)	(0.219)
	[0.018]	[0.084]	[0.627]	[0.023]
N_{tot}	3,464	3,464	3464	3464
N_{eff}	3,464	3,464	3464	3464
N_{treat}	1280	1280	1280	1280
$N_{counties}$	495	495	495	495
$\mu_{outc.}$	0.093	0.128	0.182	0.083
Covariates	Yes	Yes	Yes	Yes
Share_{sample}	0.401	0.373	0.120	0.259
Share_{US}	0.485	0.348	0.209	0.149

Note: Table displays transformed coefficient with *p<0.1; **p<0.05; ***p<0.01, followed by estimated coefficient, its standard error, and its p-value. Poisson estimation with year and adjacency county cluster fixed effects, and standard errors clustered at the county level. Set of controls described in Appendix G.

Column 1 presents effects for White people of non-Hispanic ethnicity and column 2 presents effects for all other minorities.[44] The results indicate that misreporting of officer-involved fatalities for White people are 40-54% higher in counties which allow law enforcement to certify the cause of death relative to their controls. The effects in column 2 for minorities are less clear with the fixed effect Poisson estimates showing no significant effect while the non-fixed effect Poisson show significant effects under the 10% threshold.

The following two columns further consider the two main minority groups of race-ethnicity, with column 3 showing results for Black people and column 4 showing results for people of Hispanic origin.[45] For Black people, we do not see any significant difference in the underreporting of police killings between law enforcement certifying counties and their adjacent controls. For Hispanic people, we observe 65% more police killings in law enforcement certifying counties but only in the no-fixed effect Poisson estimation. Although surprising, the absence of effects for Black people may simply reflect the low share of Black people and the lower share of underreported police killings for Black people in the analysis sample relative to the US population as a whole.

5.5.2 Underreporting in US-Mexico Border Counties

We further investigate the effects on Hispanic people. One consequence of institutions which can be exploited to circumvent accountability is that they will affect those with fewer rights. In particular, families of illegal immigrants killed by police may be less likely to bring a civil case against law enforcement. As such, law enforcement may be more likely to underreport killings of illegal immigrants. To explore, somewhat informally, this question, we use our adjacent sample approach comparing the underreporting of police killings in counties along the US-Mexico border, to their nearest inland neighbouring counties.

Table 6 presents these US-Mexico border results. To increase sample size, we include second degree adjacency counties in control counties, as presented in map A12 of Appendix D. Column 1 of Table 6 shows that counties bordering Mexico are 74% more likely to underreport police killings of people of Hispanic origin. Column 2 and 3 show that this effect is driven by lower law enforcement agency reporting of Hispanic victims to the SHR, rather than higher true police killings in the MPV data. Column 4 adds to these results showing that, for non-Hispanic people, US-Mexico bordering counties have fewer underreported police killings. This difference cannot be explained by population differences, since Mexico bordering counties display very similar average populations of

[44] The share of underreported police killings for White people and minority people do not sum up to 1 due to some deaths being classified as Unknown race.

[45] In order to ensure the race and ethnicity categories are comparable in the MPV and SHR data, we exclude Hispanic people from the Black people category by cross-checking names through a genealogy website, https://forebears.io/, and reclassifying any Black person of Spanish or Portuguese name into the Hispanic people category. We similarly reclassify Unknown race people with Spanish or Portuguese names into the Hispanic people category. The Hispanic people category will therefore also include Black race people.

Table 6: US-Mexico border effects

Dep. Var.:	Diff. L.E. homicides				
	Hisp. MPV-SHR	Hisp. SHR	Hisp. MPV	Non-Hisp. MPV-SHR	All SHR
LE certify	0.742 ***	-0.466 ***	-0.014	-0.387	0.062
	0.555	-0.628	-0.014	-0.490	0.060
	(0.164)	(0.185)	(0.139)	(0.158)	(0.150)
	[0.001]	[0.001]	[0.922]	[0.002]	[0.689]
Covariates	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes
Spec.	Pois.	Pois.	Pois.	Pois.	Pois.
N_{tot}	546	546	546	546	546
N_{eff}	532	308	539	546	448
N_{treat}	203	189	203	203	196
$N_{counties}$	76	44	77	78	64
$\mu_{outc.}$	0.400	1.013	0.970	0.332	1.605

Note: Table displays transformed coefficient with *p<0.1; **p<0.05; ***p<0.01, followed by estimated coefficient, its standard error, and its p-value. Poisson estimation with year and adjacency county cluster fixed effects, and standard errors clustered at the county level. Set of controls described in Appendix G.

Hispanic people relative to their controls and, in fact, have slightly higher populations of non-Hispanic people.

One explanation, consistent with previous results and the results of column 5, showing that there is no difference in overall reported SHR police killings, is that police agencies are aware and careful to not report excessive police killings relative to agencies in their neighbouring counties. They do, however, use discretion, possibly based on their chances of coming under scrutiny, to decide which police killings to hide from their SHR reports.

5.6 Additional Differences in Death Investigation Systems

As discussed previously, different death investigation systems and lead examiners may differ along several dimensions. As described statistically in Table A3 of Appendix F and represented in maps in Appendix D, they may be coroner or medical examiner counties, the death investigator may or may not be required to be a physician, and the coroners may be appointed or elected. Table 7 explores whether these partially overlapping differences in laws and systems lead to differences in the underreporting of police killings. Each analysis restricts the sample to adjacent treated and control counties with different death investigation systems and excludes counties in which law enforcement can certify the cause of death.

Column 1 compares the underreporting of police killings between coroner and medical examiner counties but finds no significant difference. Speaking again to the question of competence, column 2 considers differences in underreporting between counties which require a physician to conduct an autopsy and counties allowing non-physicians to perform autopsies. Although the effects are marginally significant at the 10% level, we do not find these robust to specifications which match physician and non-physician counties on urbanisation levels. Neither of these results provides convincing evidence that higher medical competence shields death investigators from intimidation and external pressure to alter their autopsy report or change the cause of death. These results are in line with the nationwide comparison of underreported police killings in medical examiner and coroner counties in GBD et al. (2021) using NVSS data.

Column 3 considers whether counties in which the death investigator is appointed rather than elected have higher police killings. Theoretically, the expected effect of the manner of selection is unclear. Elected officials may be more willing to uprightly serve their voting constituents by opposing any outside pressure to change death reports. They may, however, also want to conform to pressures by local politicians and police if they want to receive political support during elections. Ultimately, the results indicate no difference in the underreporting of police killings between elected and appointed counties. This may be because both stated effects cancel out, or simple because the election vs. appointment selection process of the death investigator is not of primary importance.

Taken together, the results from Table 7 do not indicate that our main results on underreporting depending on law enforcement death certification laws are driven by other underlying differences in death investigator characteristics or systems.

5.7 Responses to Monitoring and Threats

Previous sections have outlined the effects of different laws surrounding death investigation systems, in particular the role of certification laws with regards to law enforcement. In this section, we explore whether law enforcement agencies and agents in counties permitting them to certify the cause of death are more likely to resist the introduction of monitoring and accountability measures, in particular the use of body-worn cameras and the threat of charging police officers. The analysis further considers the effect of certification laws on homicide clearance rates, with a view to understanding whether additional scrutiny through stricter certification laws lowers police effectiveness.

We then move on to examine more generally, and not causally, whether modern-day perceived and actual threats to law enforcement are related to differences in underreporting. We first consider whether police in counties with permissive death certification laws are more likely to be assaulted or killed while on duty. Then, looking at nationwide correlations, we assess whether the permissiveness of gun and ammunition laws is related to the underreporting of police killings. In a last step, we consider whether awareness and positive or negative concern for issues raised by the BLM movement, as proxied by Google search trends and their changes, are associated with the underreporting of police

Table 7: Coroner vs Medical Examiner, Appointed vs Elected, and Physician vs Non-Physician

	Cor vs ME	Phys vs non-Phys	App vs Elec
Treatment	-0.038	0.164*	-0.006
	-0.039	0.152	-0.006
	(0.089)	(0.083)	(0.096)
	[0.663]	[0.066]	[0.950]
Covariates	Yes	Yes	Yes
FE	Yes	Yes	Yes
Spec.	Pois.	Pois.	Pois.
N_{tot}	6,223	7315	2,764
N_{eff}	5,271	6034	2,190
N_{treat}	3010	2247	356
$N_{counties}$	753	862	313
$\mu_{outc.}$	0.318	0.337	0.215

Note: Table displays transformed coefficient with *p<0.1; **p<0.05; ***p<0.01, followed by estimated coefficient, its standard error, and its p-value. (1) Coroner vs Medical Examiner adjacency sample, Treatment=Coroner, (2) Appointed vs Elected adjacency sample, Treatment=Appointed, (3) Physician vs non-Physician adjacency sample, Treatment=Physician. Poisson estimation with year and adjacency county cluster fixed effects, and standard errors clustered at the county level. Set of controls described in Appendix G.

killings.

5.7.1 Body-worn Cameras

Body-worn cameras have been proposed as an important tool to help monitor police activity. In the context of underreported police killings, body-cameras may prevent otherwise unverifiable cover-ups. Despite this promise, experimental and non-experimental results on the effects of body-worn cameras are mixed (Lum et al., 2020; Williams Jr et al., 2021). One stated problem is the apparent resistance to adopting body-worn cameras. Here, we explore whether this resistance is correlated to law enforcement death certification laws. In general, US-wide county-level information on body-worn cameras is limited. The most complete county-level source is the Law Enforcement Management and Administrative Statistics Body-Worn Camera Supplement (LEMAS-BWCS) census conducted in 2016. This census asks a random sample of law enforcement agencies when they began using body-worn cameras. It also asks questions regarding when these cameras are required to be worn.

The relation between the underreporting of police killings, whether law enforcement can certify death, the adoption and subsequent utilization of body cameras, and census response rates, can operate in many potentially endogenous ways. Laws pertaining to death certification may have, over years, contributed to cultures of impunity, rendering law enforcement agencies less likely to adopt cameras. As previously documented with the NIBRS, these law enforcement agencies may also be less willing to share detailed information on their operations. In addition, we only have a measure of whether and how many body-worn cameras were purchased, but agents may selectively choose when to turn them on. Despite these issues, which we make no claim to address adequately in this paper, we offer some first insights on how death certification laws may influence the adoption of body-cameras.

Table 8: Monitoring and Threat Effects

	LEMAS	UCR		LEOKA
Dep. Var.:	Body-cam	Clearances	Clearance rate	Assaults
	(1)	(2)	(3)	(4)
LE certify	0.005	-0.031	0.029	0.043
		-0.032		0.042
	(0.032)	(0.054)	(0.095)	(0.099)
	[0.872]	[0.556]	[0.761]	[0.668]
Covariates	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes
Spec.	Lin.	Pois.	Quasi-bin.	Pois.
N_{tot}	1,980	3464	3464	3464
N_{eff}	1,980	3219	1361	3338
N_{treat}	732	1133	544	1217
$N_{counties}$	495	460	337	477
$\mu_{outc.}$	0.217	3.392	0.565	30.525

Note: Table displays in columns 2 and 4 the transformed coefficient with *p<0.1; **p<0.05; ***p<0.01, followed by estimated coefficient, its standard error, and its p-value. Dependent variables: (1) 1 if the share of law enforcement agencies in a county claiming to operate with body-worn cameras is above the national mean of 6%, 0 otherwise, (2) Total clearances of UCR homicides, (3) Clearance rate of UCR homicides (4) Total assaults on law enforcement officers. Set of controls described in Appendix G.

Column 1 of Table 8 considers whether permissive death certification laws are correlated to the probability that law enforcement agencies within a county adopt body-worn cameras. We take as the county outcome variable for body-worn cameras an indicator equal to 1 if the share of law enforcement agencies in a county claiming to operate with body-worn cameras is above the national mean of 6%, and 0 otherwise.[46] Because the census was administered in 2016, we also drop any year thereafter from the analysis.

The results in column 1 do not suggest lenient death-certification laws influence the adoption of body-worn cameras. Although not presented in tables, we also explored heterogenous effects of law enforcement certification laws, interacted with our body-

[46] Approximately 23% of law enforcement agencies responded to the LEMAS-BWCS (only 30% were sampled) which explains the relatively low share of law enforcement agencies operating with body-worn cameras.

worn camera variable, on the underreporting of police killings, but found no significant heterogenous effects. These stated results pertain to a noisy measure of body-worn camera availability at the extensive margin. Further inquiries into the intensive use of body-worn cameras may produce additional insights.

5.7.2 Charging Officers

Speaking to another form of accountability, we also considered whether charging an officer had any effect on the underreporting of police killings in subsequent years. Unfortunately, an officer was charged in only 3.5% of true MPV police killing cases in our neighbouring county sample, and one third of those resulted in an acquittal or dropped charge(s). As a result, we only observe 18 charged or charged-convicted officers in our adjacency analysis sample over the years 2013-2019, too few for any rigorous analysis.[47]

5.7.3 Police Effectiveness

Raising the expected penalty of an officer's errors through more stringent monitoring laws may, in theory, backfire and reduce the effectiveness of police (Prendergast, 2003; Gavazza and Lizzeri, 2007). Some empirical evidence exists supporting this view (Shi, 2008; Ba and Rivera, 2019; Devi and Fryer Jr, 2020; Premkumar, 2020; Campbell, 2022). We consider here whether more lenient law enforcement certifying laws affect police performance by looking at total homicide clearances and clearance rates obtained from the UCR. Our results in Table 8 columns 2-3 do not offer any evidence that more stringent death certification laws influences police effectiveness.

5.7.4 Assaults on Police Officers

Our base independence assumption is that, prior to any changes in death investigation systems and laws in the 1960s-1980s, adjacent treatment and control counties in our analysis sample were comparable. However, after the changes in systems, different policing and reporting approaches may have affected the general hostility in a county. Homicide cover-ups by police in the late 1980s and early 1990s in counties permitting law enforcement to certify the cause of death may have incited more violent behavior towards police, which itself induced police to respond with more unwarranted lethal violence, requiring additional cover-ups of police killings.

Table 8, column 4, assesses evidence of such a cyclical process by looking at whether police in counties permitting law enforcement to certify the cause of death face more hostile environments, in the form of violent assaults. Yearly numbers of law enforcement officers who have been victims of assaults while on duty in each county are gathered from

[47] Overall in the US over that same period, we only observe 91 charged or charged-convicted officers out of 3611 police killings, or 2.5%.

the Law Enforcement Officers Killed and Assaulted (LEOKA) dataset.[48] Column 4 shows that officers in law enforcement certifying counties are no more likely to be assaulted while on duty than officers in their adjacent control counties. These results also suggest that higher lethal police action, resulting in higher unjustified killings and their cover-ups, are not a response to more hostile environments in law enforcement certifying counties.

5.7.5 Nationwide Threats and Underreporting

Departing from any causal interpretation, our data also offers a chance to describe the nationwide relation between underreported police killings and other policy relevant variables. In particular, assuming underreported police killings are a close proxy to unwarranted police killings, our data can offer a first look into the relation between unwarranted police killings and threats on police, both actual and potential.

Table 9: Threats on police and Underreporting

Dep. Var.:	MPV-SHR	MPV-SHR	SHR	MPV
	(1)	(2)	(3)	(4)
Assaults	-0.081 -0.085 (0.090) [0.347]			
Gun laws		-0.171*** -0.187 (0.031) [0.000]	0.050 0.049 (0.061) [0.423]	-0.079*** -0.082 (0.026) [0.002]
Covariates	Yes	Yes	Yes	Yes
FE	No	No	No	No
Spec.	Pois.	Pois.	Pois.	Pois.
N_{tot}	21993	21993	21993	21993
N_{eff}	21993	21993	21993	21993
N_{treat}	46757	46757	46757	46757
$N_{counties}$	3142	3142	3142	3142
$\mu_{outc.}$	0.220	0.220	0.138	0.347

Note: Table displays transformed coefficient with *p<0.1; **p<0.05; ***p<0.01, followed by estimated coefficient, its standard error, and its p-value. Poisson estimation with year fixed effects, and standard errors clustered at the county level. Set of controls described in Appendix G.

Table 9 column 1 present Poisson model estimates of underreported police killings on the arcsinh transformed value of violent assaults on police, the same variable described in the previous section. The result, controlling for our demographic, cultural, and economic covariates, do not seem to suggest that there is a US-wide correlation between these

[48]This outcome is right skewed, as shown in Appendix K. We also show maps of the averages by county over all years for these two variables.

threats on police and the underreporting of police killings.[49]

Columns 2-4 of Table 9 examine whether potential threats, proxied by state-level gun and ammunition laws and regulations, correlate with underreported police killings. In theory, operating in an environment where citizens have a higher probability of carrying a gun may give rise to premature and excessive use of lethal force. As a result, we would expect to see less underreporting of police killings in states with stricter gun regulation laws.

Our proxy for state-level gun and ammunition regulations comes from Giffords Law Center's generalized gun regulation scorecard. This is a yearly measure for each state of the strictness of gun and ammunition regulation ranging from 0-11 where 11 corresponds to a score of A and 0 to a score of F. The yearly state score weighs the many laws pertaining to ammunition and gun possession, distribution, and right-to-carry within a state. We describe and map the average state score in Appendix K.

The result in column 2 of Table 9 indicates that stricter gun and ammunition laws are correlated to lower underreporting of police killings. Columns 3 and 4 further show that this relation is due to strict gun and ammunition laws being associated to lower true police killings rather than higher reported killings. Even if only correlations, these results are concerning if we interpret the underreporting outcome as a proxy to unwarranted police killings. The correlational results may suggest that law enforcement agents respond to the potential threat of working in permissive gun law state by using excess lethal violence, resulting in unwarranted police killings which they subsequently hide.

5.7.6 Black Lives Matter and Underreporting

Another question of interest when it comes to correlation patterns is how the underreporting of police killings have been affected by the Black Lives Matter movement. A variety of measures could be used to assess the salience of the BLM movement in particular areas. Among these are the proximity to highly publicized police-killings, the scale of local BLM protests, or the general awareness and concern for issues raised by the BLM movement. The current discussion only considers this last question of awareness and concern, which we proxy by the trends in Google searches for the topic 'Black Lives Matter' in each state-year. We construct this measure by standardizing the Google trend to range from $[0, 100]$, with 100 being the highest state-year per capita search record.[50]

Taking the continuous measure of Google trends, Column 1 of Table 10 shows that, nationwide, states with higher searches for 'Black Lives Matter' topics show no correlation with the underreporting of police killing. These null results may, however, not fully describe the relationship between the concern for the BLM movement and the un-

[49] Although assaults on police do correlate with higher reported SHR and true MPV police killings.

[50] More precisely, we download each year's Google trend by state ranging from $[0, 100]$, and multiply each year-state Google trend value by the ratio of that year's Google trend searches to total searches over 2013-2019, as presented in Appendix L.

derreporting of police killings. Indeed, the Google trends measure displays a bimodal distribution as presented in Appendix L. It may be that the BLM movement only influences the underreporting of police killings beyond a certain threshold of public concern, whether favorable or unfavorable.

Table 10: Black Lives Matter and Underreporting

Dep. Var.:	MPV-SHR	MPV-SHR	MPV-SHR	MPV-SHR	MPV
	(1)	(2)	(3)	(4)	(5)
BLM_{con}	0.004 0.004 (0.005) [0.444]				
BLM_{20+}		0.269* 0.238 (0.125) [0.057]	0.090 0.086 (0.111) [0.438]	-0.217 -0.244 (0.295) [0.408]	-0.221 -0.25 (0.298) [0.401]
LE certify if $BLM_{20+} = 0$				0.395*** 0.333 (0.124) [0.007]	0.416** 0.348 (0.165) [0.035]
LE certify if $BLM_{20+} = 1$				0.806*** 0.591 (0.216) [0.006]	0.844** 0.612 (0.250) [0.014]
Covariates	Yes	Yes	Yes	Yes	Yes
FE	No	No	St	Cl	Cl+St
Spec.	Pois.	Pois.	Pois.	Pois.	Pois.
N_{tot}	21993	21993	21993	3464	3464
N_{eff}	21,433	21,433	21,433	2736	2736
N_{treat}	2526	2526	2526	930	930
$N_{counties}$	3142	3142	3142	391	391
$\mu_{outc.}$	0.22	0.22	0.22	0.289	0.289

Note: Table displays the transformed coefficient with *p<0.1; **p<0.05; ***p<0.01, followed by estimated coefficient, its standard error, and its p-value. Year fixed effects included in all specifications. *BLM* specified as continuous variable in $[0, 100]$ in column 1 and as a binary indicator equal to 1 when the continuous measure exceeds 20, and 0 otherwise in columns 2-5. Heterogeneity terms are specified as separate, not cumulative, effects depending on whether the binary *BLM* is equal to 0 or 1. Standard errors clustered at the county level. Set of controls described in Appendix G.

We explore this possibility in columns 2-5. Defining a low-high binary value for BLM concern, we find that high BLM concern in a given state[51] is associated with more underreporting. Column 3 includes state fixed effects, thereby considering the relation between changes from low-to-high or high-to-low BLM concern on changes in underreporting. There does not appear to be any relation in terms of changes. Among other explanations, the correlations from columns 2-3 may indicate that the higher concern for the BLM movement puts pressure on police to reduce visible police killings more than

[51] This is defined as a value of the Google trend above 20.

it reduces actual killings. Another possibility we cannot exclude is that states with high per capita Google searches for 'Black Lives Matter' are distinct in several dimensions affecting the underreporting of police killings.

Columns 3-4 further explore the heterogeneous effects of BLM concern in our adjacent county sample. Column 4, without state fixed effects, and column 5, with state fixed effects, indicate that counties permitting law enforcement to certify the cause of death and with high concern for the BLM movement display higher underreporting of police killings compared to the interaction effect for counties in states with low concern for the BLM movement.

These effects are consistent with previous findings of strategic manipulation of death records. Assuming Google trends are not only a proxy for the concern for the BLM movement, but also a credible measure of additional scrutiny on police, then the results suggest that police departments in counties with permissive certification laws and under additional scrutiny are even more strategic in hiding police killings than those facing less scrutiny. If instead the BLM proxy reflects searches from people opposed to the BLM movement, or result in a more negative view of the BLM movement, a negative view known by the police, then the results may indicate that law enforcement agencies take advantage of the sentiment against the BLM movement to further manipulate death records. Further investigations on this matter are important, but beyond the scope of this paper.

6 Conclusion

Rates of officer-involved fatalities in the US far surpass those in countries of similar wealth. The rate in the US is nearly 3.5 times the rate in Canada, over 25 times that in Germany, and 67 times that in England and Wales (Prison Policy, 2020). Discussions of the causes for these high rates range from the unique dangers faced by US law enforcement agents given the US's permissive gun-ownership culture, to arguments alluding to pernicious cultures of violence in police departments.

Despite these long-standing discussions, policy changes to address excess police killings in the US have been few and far between since the Supreme Court's decision in Tennessee v. Garner, 1985, regarding the use of lethal violence on a fleeing suspect. And, although the Black Lives Matter movement reignited debates around police violence, there is little evidence for reduced use of deadly force by officers, and some evidence of more underreporting.

This paper adds an important consideration to these discussions. It highlights how poorly designed institutional laws can be co-opted to hinder the accountability of state actors. It further underlines the need for a more assiduous separation of law enforcement from final assessments of the cause of death in general and in particular in cases of police killings. A first step in this process is to increase the checks and balances for death investigation offices. To date, of the 2,342 death investigation offices in the US, only 108 offices are accredited by the National Association of Medical Examiners (NAME) and only 35 are accredited by the International Association of Coroners and Medical Examiners (IACME). Part of these checks and balances should come in the form of external monitoring, ensuring that death investigators are independent from law enforcement when it comes to sensitive death investigations.

More generally, new open-source registries of police killings put together by private citizens represent an important step in holding law enforcement accountable. However, private citizens can be intimidated or threatened, so relying on individuals rather than institutions to uphold the social contract between citizens and police is unstable. While some families may seek accountability in courts based on unofficial records, police killings of people with fewer protections, such as illegal immigrants or sex workers, may never face review if omitted from official records. A second policy recommendation from this paper is therefore to allow individuals to report cases of police killings directly to the FBI, and have the FBI use their authority to investigate data discrepancies in police killings from police departments as compared to those from the public.

References

Abrams, D. S., Bertrand, M., and Mullainathan, S. (2012). Do judges vary in their treatment of race? *The Journal of Legal Studies*, 41(2):347–383.

Ang, D. (2021). The effects of police violence on inner-city students. *The Quarterly Journal of Economics*, 136(1):115–168.

Ang, D., Bencsik, P., Bruhn, J., and Derenoncourt, E. (2021). Police violence reduces civilian cooperation and engagement with law enforcement. Technical report, HKS Working Paper No. RWP21-022.

Ang, D. and Tebes, J. (2020). Civic responses to police violence. Technical report, HKS Working Paper No. RWP20-033.

Antonovics, K. and Knight, B. G. (2009). A new look at racial profiling: Evidence from the boston police department. *The Review of Economics and Statistics*, 91(1):163–177.

Anwar, S., Bayer, P., and Hjalmarsson, R. (2012). The impact of jury race in criminal trials. *The Quarterly Journal of Economics*, 127(2):1017–1055.

Arnold, D., Dobbie, W., and Yang, C. S. (2018). Racial bias in bail decisions. *The Quarterly Journal of Economics*, 133(4):1885–1932.

Ba, B. A. and Rivera, R. (2019). The effect of police oversight on crime and allegations of misconduct: Evidence from chicago. *U of Penn, Inst for Law & Econ Research Paper*, (19-42).

Baicker, K. and Jacobson, M. (2007). Finders keepers: Forfeiture laws, policing incentives, and local budgets. *Journal of Public Economics*, 91(11-12):2113–2136.

Banerjee, A., Chattopadhyay, R., Duflo, E., Keniston, D., and Singh, N. (2021). Improving police performance in rajasthan, india: Experimental evidence on incentives, managerial autonomy, and training. *American Economic Journal: Economic Policy*, 13(1):36–66.

Becker, G. S. and Stigler, G. J. (1974). Law enforcement, malfeasance, and compensation of enforcers. *The Journal of Legal Studies*, 3(1):1–18.

Benoît, J.-P. and Dubra, J. (2004). Why do good cops defend bad cops? *International Economic Review*, 45(3):787–809.

Bor, J., Venkataramani, A. S., Williams, D. R., and Tsai, A. C. (2018). Police killings and their spillover effects on the mental health of black americans: a population-based, quasi-experimental study. *The Lancet*, 392(10144):302–310.

Campbell, T. (2022). Black lives matter's effect on police lethal use of force. *Available at SSRN 3767097*.

Cheng, C. and Long, W. (2018). Improving police services: Evidence from the french quarter task force. *Journal of Public Economics*, 164:1–18.

Comey, M. L., Eng, A. R., and Pei, Z. (2022). Supercompliers. *arXiv preprint arXiv:2212.14105*.

Desmond, M., Papachristos, A. V., and Kirk, D. S. (2016). Police violence and citizen crime reporting in the black community. *American Sociological Review*, 81(5):857–876.

Devi, T. and Fryer Jr, R. G. (2020). Policing the police: The impact of "pattern-or-practice" investigations on crime. Technical report, National Bureau of Economic Research.

Dube, A., Lester, T. W., and Reich, M. (2010). Minimum wage effects across state borders: Estimates using contiguous counties. *The Review of Economics and Statistics*, 92(4):945–964.

Edwards, F., Esposito, M. H., and Lee, H. (2018). Risk of police-involved death by race/ethnicity and place, united states, 2012–2018. *American Journal of Public Health*, 108(9):1241–1248.

Edwards, F., Leeb, H., and Esposito, M. (2020). Risk of being killed by police use-of-force in the us by age. *Race/Ethnicity, and Sex, Journal of Proceedings of the National Academy of Sciences*, 116(34):1–12.

Feigenberg, B. and Miller, C. (2021). Racial divisions and criminal justice: Evidence from southern state courts. *American Economic Journal: Economic Policy*, 13(2):207–40.

Feldman, J. M., Gruskin, S., Coull, B. A., and Krieger, N. (2017). Quantifying underreporting of law-enforcement-related deaths in united states vital statistics and news-media-based data sources: A capture–recapture analysis. *PLoS medicine*, 14(10):e1002399.

Fernandez, J. M. (2019). The political economy of death: Do coroners perform as well as medical examiners in determining suicide? *Available at SSRN 3360656*.

Fryer Jr, R. G. (2019). An empirical analysis of racial differences in police use of force. *Journal of Political Economy*, 127(3):1210–1261.

Gavazza, A. and Lizzeri, A. (2007). The perils of transparency in bureaucracies. *American Economic Review*, 97(2):300–305.

GBD et al. (2021). Fatal police violence by race and state in the usa, 1980–2019: a network meta-regression. *The Lancet*, 398(10307):1239–1255.

Gelman, A., Fagan, J., and Kiss, A. (2007). An analysis of the new york city police department's "stop-and-frisk" policy in the context of claims of racial bias. *Journal of the American Statistical Association*, 102(479):813–823.

Goncalves, F. and Mello, S. (2021). A few bad apples? racial bias in policing. *American Economic Review*, 111(5):1406–1441.

Hanzlick, R. (2007). The conversion of coroner systems to medical examiner systems in the united states: a lull in the action. *The American Journal of Forensic Medicine and Pathology*, 28(4):279–283.

Hanzlick, R. L. (2014). A synoptic review of the 1954 "model postmortem examinations act". *Academic Forensic Pathology*, 4(4):451–454.

Hoekstra, M. and Sloan, C. (2022). Does race matter for police use of force? evidence from 911 calls. *American Economic Review*, 112(3):827–860.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, pages 467–475.

Knowles, J., Persico, N., and Todd, P. (2001). Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy*, 109(1):203–229.

Leflar, R. A. (1955). Drafting the model post-mortem examinations act. *ABAJ*, 41:266.

Long, W. (2019). How does oversight affect police? evidence from the police misconduct reform. *Journal of Economic Behavior & Organization*, 168:94–118.

Lum, C., Koper, C. S., Wilson, D. B., Stoltz, M., Goodier, M., Eggins, E., Higginson, A., and Mazerolle, L. (2020). Body-worn cameras' effects on police officers and citizen behavior: A systematic review. *Campbell Systematic Reviews*, 16(3).

Luzi, S. A., Melinek, J., and Oliver, W. R. (2013). Medical examiners' independence is vital for the health of the american legal system. *Academic Forensic Pathology*, 3(1):84–92.

Mastrorocco, N., Ornaghi, A., et al. (2020). *Who watches the watchmen? Local news and police behavior in the United States*. University of Warwick, Department of Economics.

Melinek, J., Thomas, L. C., Oliver, W. R., Schmunk, G. A., Weedn, V. W., and of Medical Examiners Ad Hoc Committee on Medical Examiner Independence, N. A. (2013). National association of medical examiners position paper: Medical examiner, coroner, and forensic pathologist independence. *Academic Forensic Pathology*, 3(1):93–98.

Moreno-Medina, J., Ouss, A., Bayer, P., and Ba, B. A. (2022). Officer-involved: The media language of police killings. Technical report, National Bureau of Economic Research.

Morris, K. and Shoub, K. (2023). Contested killings: The mobilizing effects of community contact with police violence. *American Political Science Review*, pages 1–17.

Nix, J., Campbell, B. A., Byers, E. H., and Alpert, G. P. (2017). A bird's eye view of civilians killed by police in 2015: Further evidence of implicit bias. *Criminology & Public Policy*, 16(1):309–340.

Olzak, S. (2021). Does protest against police violence matter? evidence from us cities, 1990 through 2019. *American Sociological Review*, 86(6):1066–1099.

Owens, E. and Ba, B. (2021). The economics of policing and public safety. *Journal of Economic Perspectives*, 35(4):3–28.

Pierson, E., Simoiu, C., Overgoor, J., Corbett-Davies, S., Jenson, D., Shoemaker, A., Ramachandran, V., Barghouty, P., Phillips, C., Shroff, R., et al. (2020). A large-scale analysis of racial disparities in police stops across the united states. *Nature Human Behaviour*, 4(7):736–745.

Prados, M. J., Baker, T., Beck, A. N., Burghart, D. B., Johnson, R. R., Klinger, D., Thomas, K., and Finch, B. K. (2022). Do sheriff-coroners underreport officer-involved homicides? *Academic Forensic Pathology*, 12(4):140–148.

Premkumar, D. (2020). Intensified scrutiny and bureaucratic effort: Evidence from policing and crime after high-profile, officer-involved fatalities. *SSRN Electron. J*, pages 1–132.

Prendergast, C. (2003). The limits of bureaucratic efficiency. *Journal of Political Economy*, 111(5):929–958.

Rad, A. N., Kirk, D. S., and Jones, W. P. (2023). Police unionism, accountability, and misconduct. *Annual Review of Criminology*, 6:181–203.

Rahman, D. (2012). But who will monitor the monitor? *American Economic Review*, 102(6):2767–2797.

Rehavi, M. M. and Starr, S. B. (2014). Racial disparity in federal criminal sentences. *Journal of Political Economy*, 122(6):1320–1354.

Ross, C. T. (2015). A multi-level bayesian analysis of racial bias in police shootings at the county-level in the united states, 2011–2014. *PloS one*, 10(11):e0141854.

Ruiz, L., Posey, B. M., Neuilly, M.-A., Stohr, M. K., and Hemmens, C. (2018). Certifying death in the united states. *Journal of Forensic Sciences*, 63(4):1138–1145.

Schwartz, G. L. and Jahn, J. L. (2020). Mapping fatal police violence across us metropolitan areas: Overall rates and racial/ethnic inequities, 2013-2017. *PloS one*, 15(6):e0229686.

Sherman, L. W. (1978). *Scandal and reform: Controlling police corruption.* Univ of California Press.

Shi, L. (2008). Does oversight reduce policing? evidence from the cincinnati police department after the april 2001 riot. *Journal of Public Economics.*

Tennenbaum, A. N. (1994). The influence of the garner decision on police use of deadly force. *J. Crim. L. & Criminology*, 85:241.

Tyler, T. R., Fagan, J., and Geller, A. (2014). Street stops and police legitimacy: Teachable moments in young urban men's legal socialization. *Journal of Empirical Legal Studies*, 11(4):751–785.

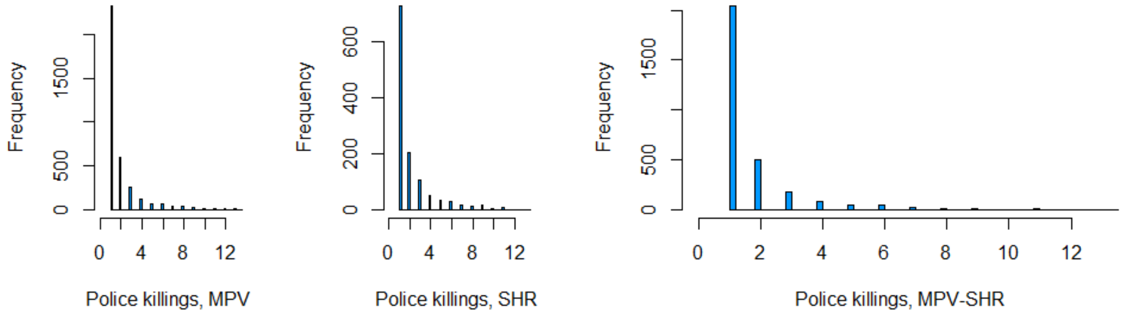
Williams Jr, M. C., Weil, N., Rasich, E. A., Ludwig, J., Chang, H., and Egrari, S. (2021). Body-worn cameras in policing: Benefits and costs. Technical report, National Bureau of Economic Research.

Zaiour, R. and Mikdash, M. (2023). The impact of police shootings on gun violence and civilian cooperation. Technical report, Available at SSRN 4390262.

Zimring, F. E. (2017). *When police kill.* Harvard University Press.

APPENDIX

A Histograms: MPV and SHR Police Killings



Note: only positive values displayed for MPV-SHR

B Maps: MPV and SHR Police Killings

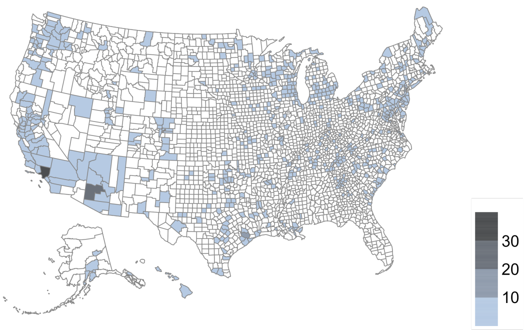


Figure A1: SHR mean police killings

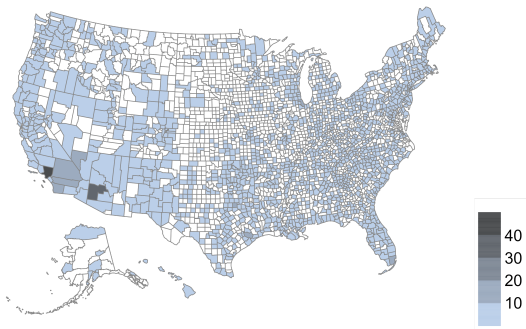


Figure A2: MPV mean police killings

C Maps: Share of Reporting Agencies to UCR and NIBRS

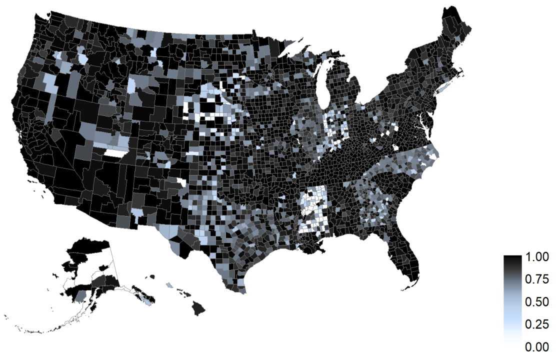


Figure A3: Share reporting UCR data

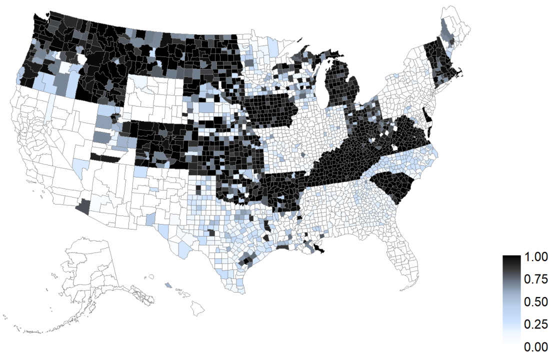


Figure A4: Share reporting NIBRS data

D Maps: Additional Differences in Death Investigation Systems

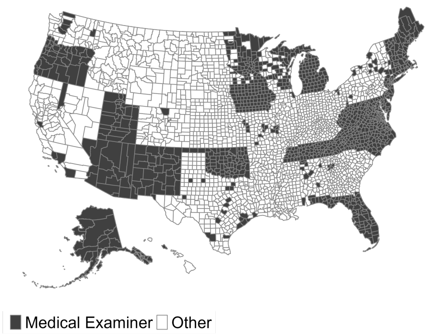


Figure A5: Medical Examiner System

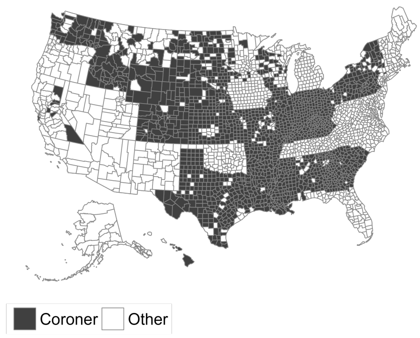


Figure A6: Coroner System

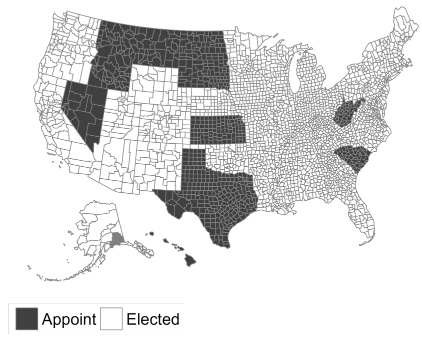


Figure A7: Appointed vs Elected

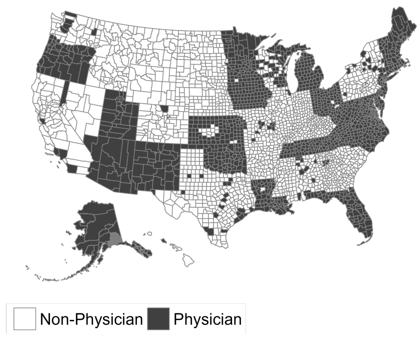


Figure A8: Physician vs non-Physician

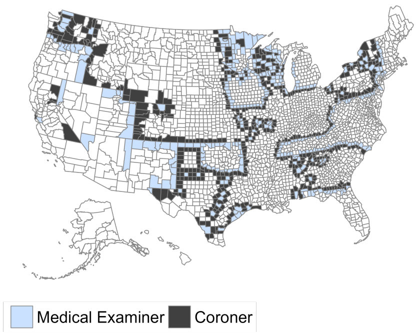


Figure A9: Coroner vs Medical Examiner analysis sample

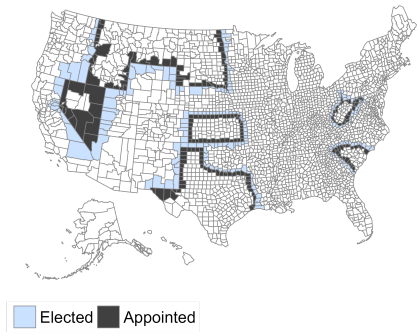


Figure A10: Appointed vs Elected analysis sample

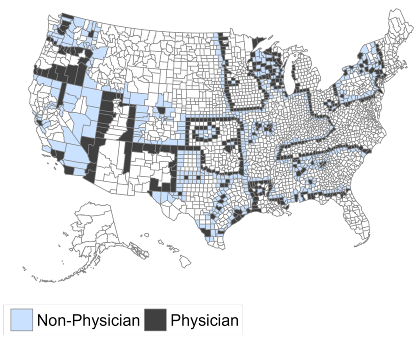


Figure A11: Physician vs non-Physician analysis sample

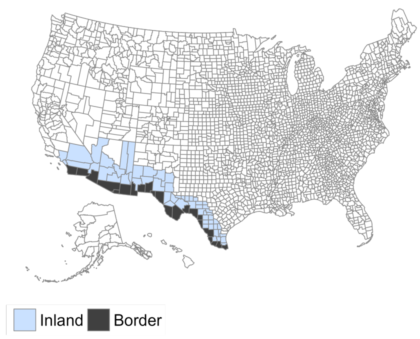


Figure A12: US-Mexico border analysis sample

E Death Investigation System: History and Current Distribution

Year	State
Before 1954	Maryland, Virginia, Vermont, Rhode Island
1961	Tennessee*
1963	Oklahoma
1964	Oregon
1967	North Carolina, Utah
1968	Maine
1969	Connecticut
1970	Delaware, Iowa*
1976	West Virginia , New Mexico
1977	Kentucky*, Alabama*
1978	New Jersey*
1979	Arkansas*
1980	Montana*
1983	Massachusetts
1986	Mississippi*, New Hampshire
1990	Georgia*
1994	North Dakota*
1996	Alaska

List of states that have implemented a state medical examiner.
* States that have implemented the state medical examiner but chose to keep a coroner or chose a medical examiner system when the mixed system was an option.
Source: Hanzlick (2014)

State	OUR DATASET	CDC Systems	CDC State ME	Hanzlick '07*
AL	ME and Cor	ME and Cor	No	-
AK	ME	ME	Yes	State ME
AZ	ME	ME	No	ME
AR	Coroner	Coroner	Yes	Coroner
CA	ME and Cor	ME and Cor	No	ME and Cor
CO*	ME and Cor	Coroner	No	ME and Cor
CT	ME	ME	Yes	State ME
DE	ME	ME	Yes	State ME
FL	ME	ME	No	ME
GA	ME and Cor	ME and Cor	Yes	ME and Cor
HI	ME and Cor	ME and Cor	No	ME and Cor
ID	Coroner	Coroner	No	Coroner
IL	ME and Cor	ME and Cor	No	ME and Cor
IN	Coroner	Coroner	No	Coroner
IA	ME	ME	Yes	State ME
KS	Coroner	Coroner	No	Coroner
KY	Coroner	Coroner	Yes	Coroner
LA	Coroner	Coroner	No	Coroner
ME	ME	ME	Yes	State ME
MD	ME	ME	Yes	State ME
MA	ME	ME	Yes	State ME
MI	ME	ME	No	ME
MN	ME and Cor	ME and Cor	No	ME and Cor
MS*	Coroner	ME and Cor	Yes	Coroner
MO	ME and Cor	ME and Cor	No	ME and Cor
MT	Coroner	Coroner	Yes	Coroner
NE	Coroner	Coroner	No	Coroner
NV**	ME and Cor	Coroner	No	Coroner
NH	ME	ME	Yes	State ME
NJ	ME	ME	Yes	State ME
NM	ME	ME	Yes	State ME
NY	ME and Cor	ME and Cor	No	ME and Cor
NC	ME	ME	Yes	State ME
ND	Coroner	Coroner	Yes	Coroner
OH	ME and Cor	ME and Cor	No	ME and Cor
OK	ME	ME	Yes	State ME
OR	ME	ME	Yes	State ME
PA	ME and Cor	ME and Cor	No	ME and Cor
RI	ME	ME	Yes	State ME
SC	Coroner	Coroner	No	Coroner
SD	Coroner	Coroner	No	Coroner
TN	ME	ME	Yes	State ME
TX	ME and Cor	ME and Cor	No	ME and Cor
UT	ME	ME	Yes	State ME
VT	ME	ME	Yes	State ME
VA	ME	ME	Yes	State ME
WA	ME and Cor	ME and Cor	No	ME and Cor
WV	ME	ME	Yes	State ME
WI	ME and Cor	ME and Cor	No	ME and Cor
WY	Coroner	Coroner	No	Coroner
DC	ME	ME	No	State ME

Note: Following our inquiries, *CO, MS and NV of our dataset differ from CDC information, NV is also different from Hanzlick (2014)

Table A1: Number of counties per Death Investigation System by State

State	Coroner	MedExam	Sheriff
AK	0	29	0
AL	57	9	1
AR	75	0	0
AZ	0	15	0
CA	7	5	46
CO	61	2	1
CT	0	8	0
DC	0	1	0
DE	0	3	0
FL	0	67	0
GA	152	5	2
HI	2	1	1
IA	0	99	0
ID	43	0	1
IL	96	1	5
IN	92	0	0
KS	92	0	13
KY	120	0	0
LA	63	0	1
MA	0	14	0
MD	0	24	0
ME	0	16	0
MI	0	83	0
MN	19	66	2
MO	96	15	4
MS	82	0	0

(a)

State	Coroner	MedExam	Sheriff
MT	19	0	37
NC	0	100	0
ND	35	0	18
NE	84	0	9
NH	0	10	0
NJ	0	21	0
NM	0	33	0
NV	0	2	15
NY	39	22	1
OH	84	2	2
OK	0	77	0
OR	0	36	0
PA	64	3	0
RI	0	5	0
SC	46	0	0
SD	55	0	11
TN	0	95	0
TX	228	21	0
UT	0	29	0
VA	0	133	0
VT	0	14	0
WA	49	6	0
WI	30	41	1
WV	0	55	0
WY	22	0	1
Total	1796	1168	172

(b)

F Summary Statistics

Table A2: Summary statistics: death investigation systems and outcomes

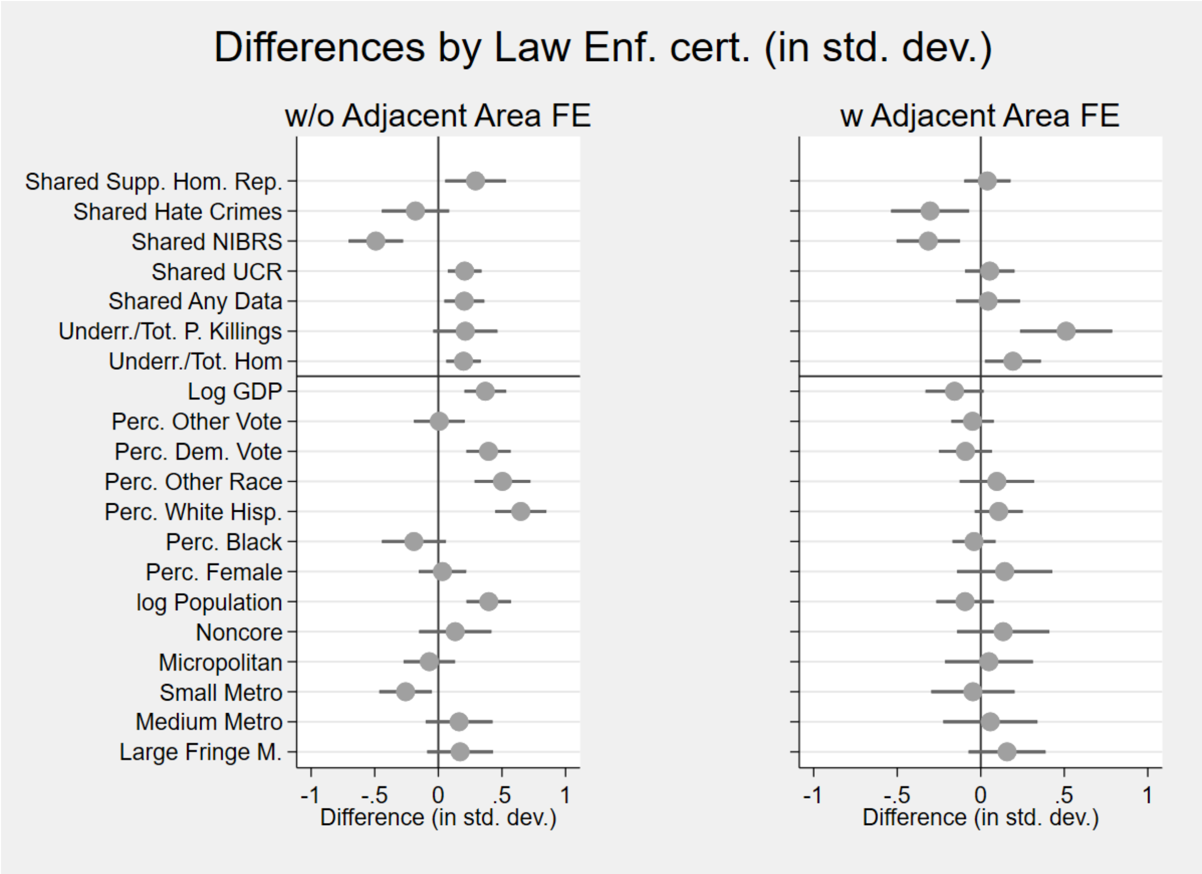
Dep. Var.:	Full population					Analysis sample	
	US	Coroner	Med.Exam.	Sh.-Coroner	LE certify	LE certify	Control
LE certify	0.11	0.07	0.05	1	1	1	0
	(0.32)	(0.25)	(0.23)	(0)	(0)	(0)	(0)
Sh.-Coroner	0.05	0	0	1	0.48	0.6	0
	(0.23)	(0)	(0)	(0)	(0.5)	(0.49)	(0)
Coroner	0.57	1	0	0	0.35	0.26	0.77
	(0.49)	(0)	(0)	(0)	(0.48)	(0.44)	(0.42)
Med.-Exam.	0.37	0	1	0	0.18	0.14	0.23
	(0.48)	(0)	(0)	(0)	(0.38)	(0.35)	(0.42)
police kill. MPV-SHR	0.21	0.13	0.32	0.3	0.28	0.35	0.15
	(0.8)	(0.48)	(1.1)	(0.88)	(0.85)	(0.95)	(0.8)
police kill. SHR	0.14	0.04	0.26	0.3	0.29	0.33	0.22
	(1.09)	(0.31)	(1.66)	(1.32)	(1.8)	(1.47)	(2.22)
police kill. MPV	0.35	0.17	0.58	0.6	0.57	0.68	0.37
	(1.54)	(0.59)	(2.27)	(1.94)	(2.33)	(2.09)	(2.88)
SHR part.	0.87	0.83	0.91	0.89	0.85	0.86	0.88
	(0.27)	(0.29)	(0.22)	(0.23)	(0.29)	(0.25)	(0.23)
NIBRS part.	0.48	0.42	0.57	0.45	0.41	0.4	0.57
	(0.46)	(0.46)	(0.46)	(0.47)	(0.45)	(0.45)	(0.45)
Tot. Homicides SHR	4.8	2.3	8.44	6.26	5.62	7.25	4.12
	(25.06)	(8.98)	(38.53)	(19.91)	(22.13)	(23.53)	(35.59)
Circ. undetermined SHR	1.73	0.72	3.17	2.56	2.23	2.94	1.02
	(10.4)	(4.78)	(15.47)	(9.27)	(10.37)	(11.58)	(8.38)
Appointed	0.22	0.29	0.07	0.56	0.32	0.34	0.47
	(0.42)	(0.45)	(0.25)	(0.5)	(0.47)	(0.48)	(0.5)
Physician	0.45	0.14	1	0	0.18	0.14	0.39
	(0.5)	(0.35)	(0)	(0)	(0.38)	(0.35)	(0.49)

Table A3: Summary statistics: covariates

Dep. Var.:	Full population					Analysis sample	
	US	Coroner	Med.Exam.	Sh.-Coroner	LE certify	LE certify	Control
Large central Metro	0.02	0	0.05	0.02	0.02	0.02	0.03
	(0.15)	(0.06)	(0.22)	(0.13)	(0.13)	(0.15)	(0.16)
Large fringe Metro	0.12	0.1	0.16	0.06	0.04	0.08	0.05
	(0.32)	(0.3)	(0.36)	(0.23)	(0.21)	(0.27)	(0.23)
Medium Metro	0.12	0.1	0.14	0.09	0.09	0.1	0.08
	(0.32)	(0.31)	(0.35)	(0.29)	(0.28)	(0.31)	(0.28)
Micropolitan	0.2	0.22	0.18	0.17	0.21	0.17	0.23
	(0.4)	(0.42)	(0.38)	(0.38)	(0.4)	(0.38)	(0.42)
Non-core	0.42	0.46	0.35	0.55	0.55	0.54	0.53
	(0.49)	(0.5)	(0.48)	(0.5)	(0.5)	(0.5)	(0.5)
Small Metro	0.11	0.11	0.12	0.1	0.1	0.09	0.08
	(0.32)	(0.31)	(0.32)	(0.31)	(0.3)	(0.29)	(0.27)
Population (per 10,000)	10.27	5.43	17.4	12.66	11.87	15.75	10.55
	(32.9)	(10.65)	(49.28)	(38.79)	(38.95)	(42.83)	(62.15)
Fem. perc	0.5	0.5	0.5	0.49	0.49	0.49	0.49
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
White (non-Hisp.) perc.	0.77	0.77	0.76	0.76	0.76	0.76	0.81
	(0.2)	(0.2)	(0.2)	(0.22)	(0.22)	(0.23)	(0.2)
Black perc.	0.09	0.1	0.09	0.02	0.02	0.02	0.03
	(0.14)	(0.16)	(0.12)	(0.05)	(0.04)	(0.05)	(0.08)
Hisp. perc.	0.09	0.09	0.08	0.12	0.13	0.12	0.09
	(0.13)	(0.14)	(0.11)	(0.15)	(0.17)	(0.16)	(0.14)
Asian-PacIsl. perc.	0.02	0.01	0.02	0.03	0.02	0.02	0.02
	(0.03)	(0.02)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)
Native perc.	0.02	0.02	0.03	0.05	0.05	0.05	0.04
	(0.08)	(0.06)	(0.09)	(0.12)	(0.12)	(0.13)	(0.11)
REP. Perc. votes	0.62	0.65	0.57	0.62	0.63	0.62	0.64
	(0.15)	(0.14)	(0.15)	(0.17)	(0.18)	(0.18)	(0.15)
DEM. Perc. votes	0.35	0.32	0.39	0.33	0.32	0.33	0.31
	(0.15)	(0.15)	(0.15)	(0.16)	(0.17)	(0.17)	(0.15)
Other Perc. votes	0.04	0.04	0.04	0.05	0.05	0.05	0.05
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
GDP (per 100,000)	56.31	25.39	105.26	61.9	61.09	81.49	68.98
	(242.59)	(67.71)	(380.09)	(212.87)	(217.84)	(246.81)	(426.14)

G Balancing Tables

Figure A13: Balance Tables



H Robustness Results

Table A4: Robustness checks of authorising law enforcement to certify cause of death

Dep. Var.:	Diff. L.E. homicides										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
LE certify	0.629**	0.436***	0.461***	0.493***	0.445***	0.528***	0.799***	0.404***	0.751***	0.510***	0.616***
	0.488	0.362	0.379	0.401	0.368	0.424	0.587	0.339	0.560	0.412	0.48
	(0.216)	(0.123)	(0.123)	(0.122)	(0.124)	(0.120)	(0.189)	(0.123)	(0.162)	(0.094)	(0.148)
	[0.024]	[0.003]	[0.002]	[0.001]	[0.003]	[0.000]	[0.002]	[0.006]	[0.001]	[0.000]	[0.001]
All covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spec.	Pois.	Pois.	Pois.	Pois.	Pois.	Pois.	Pois.	Pois.	Pois.	Pois.	Pois.
Offset	No	No	No	No	No	No	Yes	Yes	No	No	No
N_{tot}	3464	3464	3,464	3,464	3464	3464	912	355	2219	6551	6551
N_{eff}	1771	2736	2,736	2,736	2736	2736	822	329	1281	6173	6173
N_{treat}	574	930	930	930	930	930	450	232	476	1707	1707
$N_{counties}$	253	391	391	391	391	391	184	93	183	882	882
$\mu_{outc.}$	0.349	0.289	0.289	0.289	0.280	0.262	0.105	0.622	0.293	0.223	0.223

Note: Table displays transformed coefficient with *p<0.1; **p<0.05; ***p<0.01, followed by estimated coefficient, its standard error, and its p-value. (1) includes adjacent county cluster fixed effects and state fixed effects, (2) includes a covariate indicating whether death investigator was a Physician (which is a potential bad control) (3) drops the potential bad control total homicide variable, (4) adds a covariate capturing the stringency of gun and ammunition laws by state, (5) Outcome is MPV police killings - MPV off-duty police killings - SHR police killings, (6) Outcome is MPV police killings by firearm (+ other) - SHR police killings, (7) offsets on the total homicide variable instead of controlling for it, (8) offsets on the total MPV police killings, (9) matches treated and control counties with the same urbanisation level, (10) allows for a second degree adjacency counties to be included, (11) extrapolated effects weighting by propensity to be in sample. Poisson estimation with year fixed effects and standard errors clustered at the county level. Set of controls described in Appendix G.

I Identification of Effect Decomposition

Table A5: Extended County Types

Type	W^1	W^0	Y^1	Y^0
$\mathbf{a^W a^Y}$	**1**	**1**	**1**	**1**
$\mathbf{a^W n^Y}$	**1**	**1**	**0**	**0**
$\mathbf{a^W c^Y}$	**1**	**1**	**1**	**0**
$a^W d^Y$	1	1	0	1
$\mathbf{n^W a^Y}$	**0**	**0**	**1**	**1**
$\mathbf{n^W n^Y}$	**0**	**0**	**0**	**0**
$\mathbf{n^W c^Y}$	**0**	**0**	**1**	**0**
$n^W d^Y$	0	0	0	1
$\mathbf{c^W a^Y}$	**1**	**0**	**1**	**1**
$\mathbf{c^W n^Y}$	**1**	**0**	**0**	**0**
$\mathbf{c^W c^Y}$	**1**	**0**	**1**	**0**
$c^W d^Y$	1	0	0	1
$d^W a^Y$	0	1	1	1
$d^W n^Y$	0	1	0	0
$d^W c^Y$	0	1	1	0
$d^W d^Y$	0	1	0	1

To reduce notational burden, we exclude in the proof the subscript i. Assumption A.II fixes all probabilities of non-bold types in Table A5 to 0. As a result, and using assumption A.I, we can derive the shares presented in equations 2. Simple algebra shows that under the additional Assumption A.III, with $\Pr(c^W n^Y) = 0$ and $\Pr(n^W c^Y) = 0$ all remaining joint probabilities in equation 3 are identified. Conditional probabilities are immediately obtained applying Bayes' rule.

J Estimation and Additional Results of Effect Decomposition

Table 3 presents robust estimates of complier effects accounting for adjacency cluster fixed effects. These linear specifications are, however, not best suited to estimate type probabilities for probability types with no compliers. Table A6 presents a full set of results when calculating each probability type from its sample mean with bootstrap p-values presented in brackets. Based on our potential outcome model, we find that $\Pr(a^Y) = 8.6\%$ of counties show underreporting of police killings regardless of the leniency of certification laws. An additional $\Pr(c^Y) = 9.3\%$ of counties underreport only when facing permissive certification laws. The majority of counties, $\Pr(n^Y) = 82.1\%$, never underreport police killings. We also find that $\Pr(c^W) = 9.2\%$ of counties report excess undetermined homicides only when facing permissive certification laws.

We further find that $\Pr(c^Y) = 9.3$ and $\Pr(c^W c^Y) = 9.2$ are slightly higher than those estimated in our more robust main specification. They still, however, convey the same conclusion, namely that the difference in underreporting between law enforcement certifying and non-certifying counties is largely driven by counties in which law enforcement underreport police killings and simultaneously over-report excess undetermined homicides. From Table A6, we can also see that, conditional on underreporting police killings, a county will almost surely report excess undetermined homicides, $\Pr(c^W|c^Y) = 99.3$. Similarly, conditional on reporting excess undetermined homicides, a county will almost surely underreport police killings, $\Pr(c^Y|c^W) = 99.7$.

Finally, we also notice that a small share of counties, $\Pr(a^W a^Y) = 4.4\%$, may always be reclassifying underreported police killings in the circumstances undetermined category regardless of death certification laws. This last effect does not depend on assumption A.III for identification. In contrast, if assumption A.III were incorrect, allowing $\Pr(c^W n^Y) > 0$ and $\Pr(n^W c^Y) > 0$, then $\Pr(c^W|c^Y)$ and $\Pr(c^Y|c^W)$ would be upper bounds and $\Pr(c^W a^Y)$ would be a lower bound. Since $\Pr(c^W|c^Y) \approx 1$, $\Pr(c^Y|c^W) \approx 1$ and $\Pr(c^W a^Y) \approx 0$, these estimated bounds would be uninformative.

Table A6: Re-classification Type Probabilities

$a^W a^Y$	0.044 [0.000]	$c^W a^Y$	0.000 [0.526]	$n^W a^Y$	0.041 [0.000]	a^Y	0.086 [0.000]	$c^W\|a^Y$	0.003 [0.526]
$a^W c^Y$	0.001 [0.419]	$c^W c^Y$	0.092 [0.000]	$n^W c^Y$	0.000 [-]	c^Y	0.093 [0.000]	$c^W\|c^Y$	0.993 [0.000]
$a^W n^Y$	0.400 [0.000]	$c^W n^Y$	0.000 [-]	$n^W n^Y$	0.421 [0.000]	n^Y	0.821 [0.000]	$c^W\|n^Y$	0.000 [-]
a^W	0.445 [0.000]	c^W	0.092 [0.000]	n^W	0.462 [0.000]				
$c^Y\|a^W$	0.001 [0.419]	$c^Y\|c^W$	0.997 [0.000]	$c^Y\|n^W$	0.000 [-]				

Note: Table displays sample means with bootstrap p-values in brackets. $N_{tot} = 3,464$, $N_{counties} = 495$. Bootstrap p-values estimated as share of 399 bootstrap estimates smaller or equal to 0.

The results presented in Table A6 are calculated from empirical means. One problem we must address when calculating these means is that some ratios $\underline{W}_i$ are undefined if no UCR total homicides are observed in a given county-year. Such instances are non-negligible in the data. They appear in 61% of county-year observations. A priori, it is unclear how to redefine W_i in these cases. We found that dropping these variables changes importantly the estimates of $\Pr(a^Y)$, $\Pr(c^Y)$ and $\Pr(n^Y)$, none of which depend on W_i in the estimation. As a result, in our main results, we opt to randomly impute a value of 0 or 1 to W_i generated from a binomial distribution. In effect, this implies that these undefined instances do not contribute to any effects. A last option we can consider is imputing values of 0 to all undefined instances of W_i. We present results for means with this imputation in Table A7. Unsurprisingly, this imputation heavily influences the marginal probabilities $\Pr(a^W)$, $\Pr(c^W)$ and $\Pr(n^W)$. We also find it changes the estimates of $\Pr(a^W n^Y)$, $\Pr(n^W n^Y)$, and $\Pr(c^W|a^Y)$. In general, this imputation leads to less robust estimates of probabilities which lie below 0 or above 1.

Table A7: Re-classification Type Probabilities: imputing 0 to undefined W_i

$a^W a^Y$	0.038	$c^W a^Y$	-0.001	$n^W a^Y$	0.049	a^Y	0.086	$c^W\|a^Y$	-0.013
	[0.000]		[0.584]		[0.000]		[0.000]		[0.584]
$a^W c^Y$	-0.027	$c^W c^Y$	0.120	$n^W c^Y$	0.000	c^Y	0.093	$c^W\|c^Y$	1.294
	[0.992]		[0.000]		[0.000]		[0.000]		[0.000]
$a^W n^Y$	0.116	$c^W n^Y$	0.000	$n^W n^Y$	0.705	n^Y	0.821	$c^W\|n^Y$	0.000
	[0.000]		[-]		[0.000]		[0.000]		[-]
a^W	0.126	c^W	0.119	n^W	0.755				
	[0.000]		[0.000]		[0.000]				
$c^Y\|a^W$	-0.216	$c^Y\|c^W$	1.010	$c^Y\|n^W$	0.000				
	[0.992]		[0.000]		[-]				

Note: Table displays sample means with bootstrap p-values in brackets. $N_{tot} = 3,464$, $N_{counties} = 495$. Bootstrap p-values estimated as share of 399 bootstrap estimates smaller or equal to 0.

Alternative estimation approaches which control for observed covariates may be more efficient. In the simplest specifications, one may opt for specifying type probabilities using logit specifications instead of empirical means. Tables A8 and A9 present results for type probabilities specified by logit functions. More precisely, we first estimate three models: a logit estimation of Y on D, a logit estimation of W on D conditioning on the sample with $Y = 0$, and a logit estimation of W on D conditioning on the sample with $Y = 1$.[52] We then generate bootstrap values of each type probability estimate and present the share of these smaller or equal to zero as bootstrap p-values.

Table A8 presents results from logit specifications when not including any covariates which allows us to inspect how much functional form affects our estimates. In general, we find the estimated effects close to those in our main table using empirical means. Table A9 adds our full set of covariates to the logit estimations. We notice that $\Pr(c^W)$ reduces in size by about two thirds and $\Pr(c^Y)$ reduces by about one third. These estimates seem to be at odds with all our previously discussed estimates as well as those from a linear model (not presented). In contrast, all other estimates of joint probabilities seem reasonably close. Because of the change in marginal probabilities for compliers, the estimated conditional probabilities for different types blow up to values larger than 1 or lower than 0.

[52] Several other variations produced (even) less robust output than those presented, i.e. probability estimates above 1 or below 0.

Table A8: Re-classification Type Probabilities: Logit no-covariates

$a^W a^Y$	0.049	$c^W a^Y$	-0.006	$n^W a^Y$	0.043	a^Y	0.086	$c^W \| a^Y$	-0.068
	[0.000]		[1.000]		[0.000]		[0.000]		[1.000]
$a^W c^Y$	-0.001	$c^W c^Y$	0.094	$n^W c^Y$	0.000	c^Y	0.093	$c^W \| c^Y$	1.010
	[1.000]		[0.000]		[0.000]		[0.000]		[0.000]
$a^W n^Y$	0.415	$c^W n^Y$	0.000	$n^W n^Y$	0.406	n^Y	0.821	$c^W \| n^Y$	0.000
	[0.000]		[-]		[0.000]		[0.000]		[-]
a^W	0.463	c^W	0.088	n^W	0.449				
	[0.000]		[0.000]		[0.000]				
$c^Y \| a^W$	-0.002	$c^Y \| c^W$	1.067	$c^Y \| n^W$	0.000				
	[1.000]		[0.000]		[-]				

Note: Table displays probabilities generated from fitted values of logit estimations with bootstrap p-values in brackets. $N_{tot} = 3,464$, $N_{counties} = 495$. Bootstrap p-values estimated as share of 399 bootstrap estimates smaller or equal to 0.

Table A9: Re-classification Type Probabilities: Logit with covariates

$a^W a^Y$	0.050	$c^W a^Y$	-0.004	$n^W a^Y$	0.060	a^Y	0.106	$c^W \| a^Y$	-0.039
	[0.000]		[1.000]		[0.000]		[0.000]		[1.000]
$a^W c^Y$	-0.033	$c^W c^Y$	0.066	$n^W c^Y$	0.000	c^Y	0.034	$c^W \| c^Y$	1.067
	[1.000]		[0.000]		[0.000]		[0.000]		[0.000]
$a^W n^Y$	0.446	$c^W n^Y$	0.000	$n^W n^Y$	0.414	n^Y	0.860	$c^W \| n^Y$	0.000
	[0.000]		[-]		[0.000]		[0.000]		[-]
a^W	0.463	c^W	0.062	n^W	0.475				
	[0.000]		[0.000]		[0.000]				
$c^Y \| a^W$	-0.070	$c^Y \| c^W$	1.067	$c^Y \| n^W$	0.000				
	[1.000]		[0.000]		[-]				

Note: Table displays probabilities generated from fitted values of logit estimations with bootstrap p-values in brackets. $N_{tot} = 3,464$, $N_{counties} = 495$. Bootstrap p-values estimated as share of 399 bootstrap estimates smaller or equal to 0.

K Histograms and Yearly Mean Maps: Threats to Police

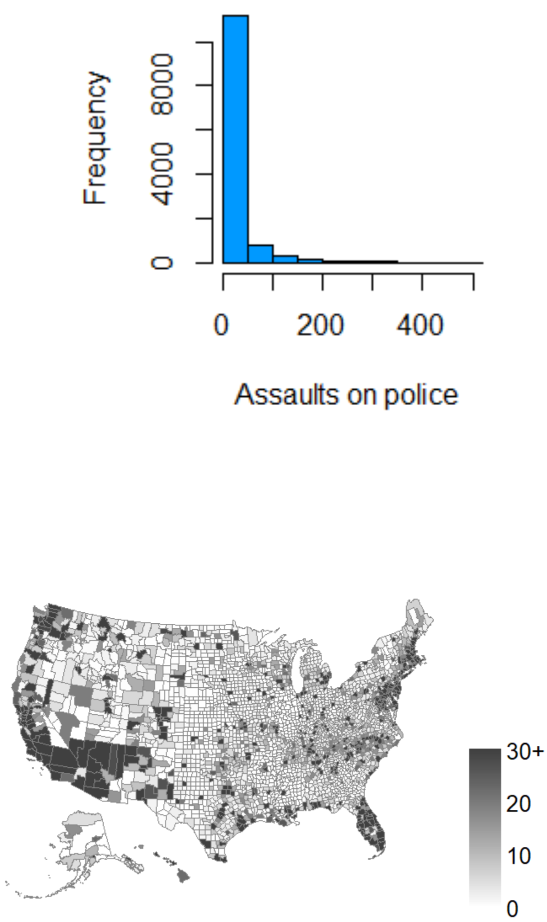


Figure A14: Assaults on police

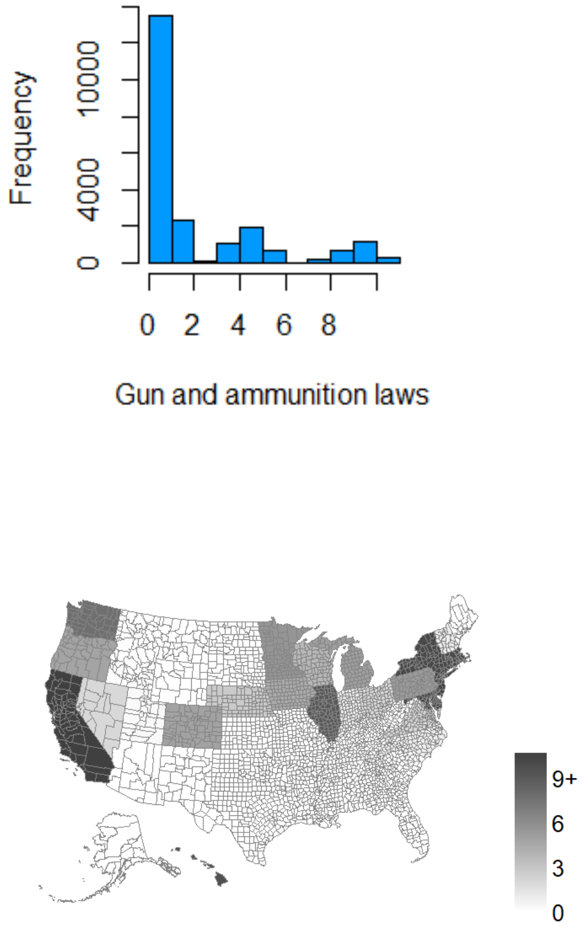
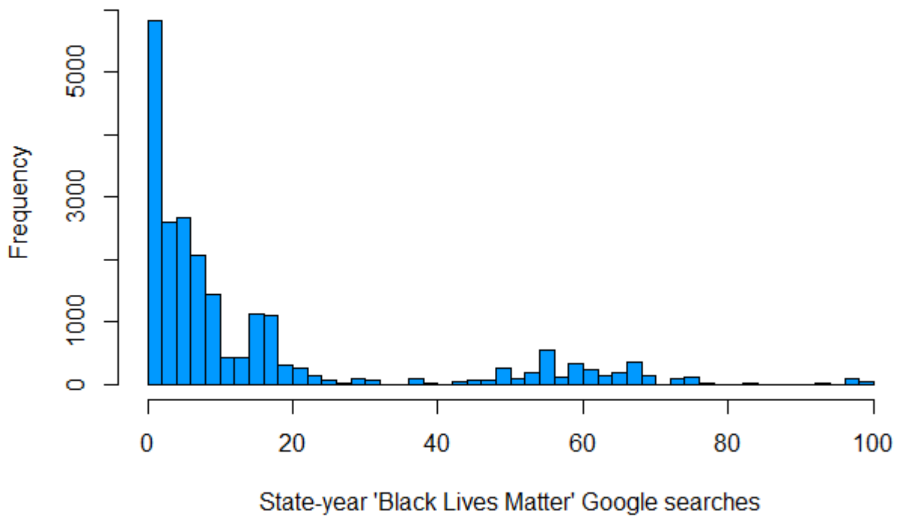


Figure A15: Gun regulations (11=A, 0=F)

L Black Lives Matter Google Trends



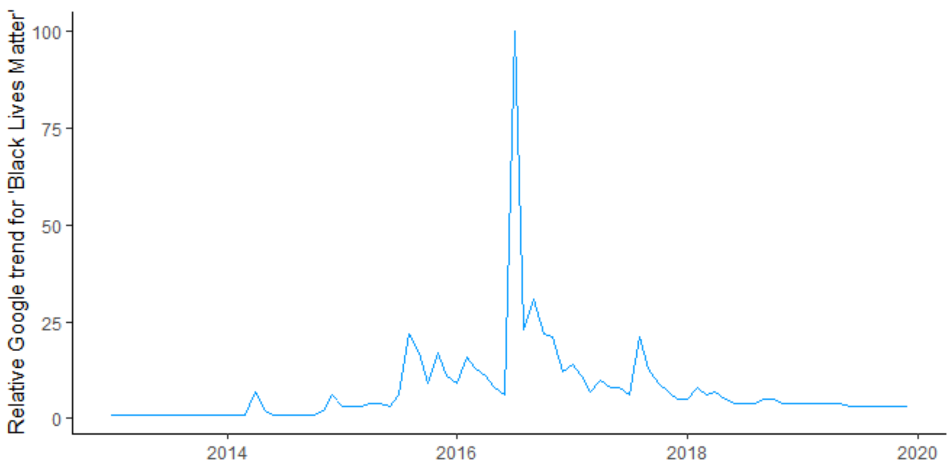


Table A10: Description of Variables

Outcome Variables	**Description**
map_kill_pol	The killing of a felon by a law enforcement officer in the line of duty. This is the Official number of police killings reported by the FBI .
mpv_kill_pol	Total incidents where a law enforcement officer (off- duty or on-duty) applies, on a civilian, lethal force resulting in the civilian being killed whether it is considered “justified” or “unjustified” by the US Criminal Legal System. We consider MPVthis to be our True value of police killings.
diff_kill_pol	Difference between mpv_kill_pol and map_kill_pol. When the difference is negative, namely when the official killings are more than the true police killings, the variable is set to zero
diff_kill_black	Difference between mpv_kill_black and map_kill_black. When the difference is negative, namely when the official killings are more than the true police killings, the variable is set to zero
diff_kill_hisp	Difference between mpv_kill_hisp and map_kill_hisp. When the difference is negative, namely when the official killings are more than the true police killings, the variable is set to zero. The variable includes Black people and White people with Hispanic origins. Black’s Hispanic origin has been retrieved with Forebears analyzing their names and accounting for a percentage of Hisponosphere above 50%. While the information for White Hispanic people is already embedded in the datasets.
diff_kill_white	Difference between mpv_kill_white and map_kill_white. When the difference is negative, namely when the official killings are more than the true police killings, the variable was set to zero
diff_kill_nwhite	Difference between mpv_kill_nwhite and map_kill_nwhite, namely all the Non-White categories of race and ethnicity
diff_kill_nblack	Difference between mpv_kill_nblack and map_kill_nblack, namely all the Non-Black categories of race and ethnicity
map_hom_tot	Total Homicides on the murder accountability project dataset which is based on SHR
mental_illness	A Person who died after an interaction with the police was allegedly showing mental health issues, as registered in the Mapping Police Violence database.

Table A10 continued from previous page

	Description
circumstance_undetermined	Person died after an interaction with the police under circumstances undetermined as reported in the Supplementary Homicide Report (SHR).
lemas_body_cam	law enforcement agencies in a county claiming to operate with body-worn cameras, Body-Worn Camera Supplement (LEMAS-BWCS) census conducted in 2016
leoka_homicides	Total homicides of law enforcement officers from yearly numbers of law enforcement officers who have been victims of felony homicides while on duty in each county are gathered from the Law Enforcement Officers Killed and Assaulted, LEOKA
leoka_assault	Total homicides of law enforcement officers from yearly numbers of law enforcement officers who have been victims of felony assaults while on duty in each county are gathered from the Law Enforcement Officers Killed and Assaulted, LEOKA
UCR	Agencies participate voluntarily and submit their crime data either through a state UCR program or directly to the FBI's UCR Program. cde.ucr.cjis.gov/LATEST
NIBRS	National Incident-Based Reporting System (NIBRS) Participation Status Indicates if the agency has submitted NIBRS data. Uniform Crime Reporting (UCR) Program - NIBRS Participation by State cde.ucr.cjis.gov/LATEST
Treatments	
LE_cert	Law Enforcement is allowed to certify the cause of death in the counties where it takes value one, zero otherwise.
SherCoroner*	Takes value one when this is the Death Investigation System implemented in the county of Interest. Namely, the Sheriff-Coroner is the person entitled to determine the cause of death of a deceased person
Coroner*	Takes value one when this is the Death Investigation System implemented in the county of Interest. Namely the Coroner is the person entitled to determine the cause of death of a deceased person
MedExaminer*	Takes value one when this is the Death Investigation System implemented in the county of Interest. Namely the Medical Examiner is the person entitled to determine the cause of death of a deceased person
Physician	Counties where the death investigator is required to be a physician.
Appointed	Counties where the death investigator is appointed and not elected.
Covariates	

Table A10 continued from previous page

	Description
Large Central Metro	Large central metro counties are counties in metropolitan statistical areas (MSA) of 1 million or more population. Urbanization size of the county according to the 2013 urbanization. www.cdc.gov
Large Fringe Metro	Large fringe metro counties are counties in MSAs of 1 million or more population that do not qualify as large central. Urbanization size of the county according to the 2013 urbanization. www.cdc.gov
Medium Metro	Medium metro counties are counties in MSAs of 250,000 to 999,999 population. Urbanization size of the county according to the 2013 urbanization. www.cdc.gov
Small Metro	Small metro counties are counties in MSAs of less than 250,000 population. Urbanization size of the county according to the 2013 urbanization.www.cdc.gov
Microp	Micropolitan counties are counties in micropolitan statistical areas. Urbanization size of the county according to the 2013 urbanization. www.cdc.gov
Noncore	Non-core counties are nonmetropolitan counties that are not in a micropolitan statistical area. Urbanization size of the county according to the 2013 urbanization. www.cdc.gov
Population	Population of the selected county in a determined year, data from census www.census.gov
fem_perc	Percentage of the female population according the census data www.census.gov
Black_perc	Percentage of the Black population according the census data www.census.gov
WHisp_perc	Percentage of the White Hispanic population according the census data www.census.gov
WNonHisp_perc	Percentage of the White population according the census data www.census.gov
Other_perc	Percentage of the Asian-Pacific Islanders and Native American population according to the census data www.census.gov
perc_votes.REP	Voting percentages during the Mayor elections in counties' major cities - Republican party, Presidential election data
perc_votes.DEM	Voting percentages during the Mayor elections in counties' major cities - Democratic party,

Table A10 continued from previous page

	Description
	Presidential election data
perc_votes.OTHER	Voting percentages during the Mayor elections in counties' major cities - Other parties, Presidential election data
GDP	Gross Domestic Product levels per county during the seven-year period. Data from Bureau of Economic Analysis Bureau of Economic Analysis [CAGDP1:GDP Summary by County and MSA]

*Note:**The variables have been created following the Census of Medical Examiners' and Coroners' Offices (MECO) Series published in December 2021. The missing values have been added one by one by searching through the CDC website (www.cdc.gov) and for each state indicating a mixed system we have gone through all the counties' websites looking whether they had a Coroner, Sheriff-Coroner or a Medical Examiner.