

# Black Lives Matter's Effect on Police Lethal Use-of-Force\*

Travis Campbell<sup>†</sup>

May 13, 2021

## Abstract

Has Black Lives Matter influenced police lethal use-of-force? A difference-in-differences design finds census places with protests experienced a 15% to 20% decrease in police homicides from 2014 through 2019, around 300 fewer deaths. This decrease was prominent when protests were large or frequent. Potential mechanisms behind the reduction include police agencies obtaining body-worn cameras to curtail force and depolicing following a so-called ‘Ferguson Effect.’ Fewer property crimes, but more murders, were reported to agencies with local protests; in contrast, the property crime clearance rate fell. Demographic imbalance by protest status and limited variation in treatment timing warrant a cautious interpretation.

**JEL Codes:** K42, Z13, D74

**Keywords:** Black Lives Matter, Police, Law and Economics, Police Homicides, Use-of-force, Law Enforcement, Ferguson Effect, Depolicing

---

\*I am grateful to Lee Badgett, Deepankar Basu, Arin Dube, Nancy Folbre, Ina Ganguli, Kathryne Young, and the participants of both the 2021 Allied Social Science Associations conference and the 2019 UMass/New School graduate student workshop for helpful feedback on earlier drafts of this paper. This paper also benefited immensely from correspondence with Vladimir Kogan. All errors remain my own.

<sup>†</sup>Department of Economics, University of Massachusetts Amherst, 105 Gordon Hall, 412 North Pleasant Street, Amherst, MA 01002. For correspondence, email: [tbcampbell@umass.edu](mailto:tbcampbell@umass.edu).

## I. Introduction

Reacting to the acquittal of George Zimmerman for the killing of Trayvon Martin in 2013, Alicia Garza posted her reaction to Facebook:

“black people. I love you. I love us. Our lives matter.”

This post inspired activist Patrisse Cullors to create a viral Twitter tag #blacklivesmatters and, with the help of activist Opal Tometi, Black Lives Matter (BLM) was born. In the wake of the 2014 police killings of Eric Garner in New York City and Michael Brown in Ferguson, Missouri, BLM transformed into the massive protests movement it is known as today. BLM did not merely make headlines, it made change, generating publicity and political pressure that led to significant policing reforms, leading the U.S. Department of Justice to distribute 21,000 policy body cameras to local law enforcement and to force eight cities to consent to improvements in police practices. Local actions may have had even greater effects than national publicity. Many of the direct and indirect effects of BLM are impossible to quantify, but a variety of data sources make it possible to compare police homicides in locations where protests took place with locations where they did not.

Has BLM altered police lethal use-of-force? This study interrogates this question using nonprofit data on police killings from Fatal Encounters Dot Org, published data on BLM protests from [Trump et al. \(2018\)](#) during 2014-2015, and web scraped data from 2015-2019 from Ainsley’s database of BLM protests. Stacked difference-in-differences (DID) estimates suggest that places with BLM protests had 15% to 20% fewer incidents of lethal use-of-force than places without BLM protests, approximately 300 fewer police homicides.

There are several potential challenges in estimating the effect of BLM protests on police homicides. First, there is currently no federal database with credible data on police killings, a long-standing problem ([Fyfe, 2002](#); [Banks et al., 2015](#); [Klinger et al., 2016](#); [White, 2016](#)). Second, a city’s characteristics may influence the likelihood of both BLM protests and police lethal force, some of which may be unobservable. Third, since BLM protests are motivated by police killings, police homicides likely rise before BLM protests, but only in cities where the protests occur ([Trump et al., 2018](#); [Skoy, 2020](#)). This pre-trend difference would break the identifying assumption of DID estimators used in some related research ([Mazumder, 2019](#); [Cunningham and Gillezeau, 2018](#)). Forth, the prevalent rainfall instrumental variable design, which uses rainfall to isolate exogenous variation in protest participation, is not valid.

I address the above problems with a stacked DID design that leverages variation in BLM protests’ location and timing to uncover the BLM’s effect, contrasting four different

estimators that address the above issues.<sup>1</sup> The benchmark specification is an unweighted two-way fixed effects estimator. The second estimator is per capita population-weighted least squares regressions, which accounts for any population-driven variance from the media neglecting protests or police homicides in less populated areas (Madestam et al., 2013).<sup>2</sup> The third estimator allows for semi-parametric selection on pre-protest correlates of use-of-force or BLM protests to address concerns with confounding variables.<sup>3</sup> To inoculate against parallel trend violations, the fourth estimator is synthetic DID, which balances police homicides between treated and controls places before protests and between control places before and after demonstrations initiate elsewhere (Arkhangelsky et al., 2019).<sup>4</sup>

Protests do not likely alter lethal use-of-force directly; rather, the change may stem from indirect channels like use-of-force regulations (Terrill and Paoline, 2017; McElvain and Kposowa, 2008), body-worn cameras (Ariel et al., 2015), officer demographics (Ba et al., 2021; Ridgeway, 2020; Tregle et al., 2019; Hoekstra and Sloan, 2020), and police disengagement from active work, so-called *depolicing* (Devi and Fryer Jr, 2020; Premkumar, 2020; Kochel, 2019; Cheng and Long, 2018). I use a stacked DID to gauge the potential of these mechanisms. I find that agencies with local protests become more likely to obtain body-cameras and expand community policing but forego some black officer employment and officer experience. I also present qualitative evidence to support the body camera finding; agencies with local protests are more likely (less likely) to self-report that they obtained body-worn cameras to reduce use-of-force and improve community perception (improve evidence quality and reduce agency liability). The fall in lethal force may also stem from depolicing if protests create a so-called *Ferguson Effect*; the idea that when public attention is drawn to police scandals, police morale falls, leading to a simultaneous rise in criminal offenses and decline in low-level arrests (Premkumar, 2020). I find that fewer property crimes, but more murders, are reported to agencies with local protests, while the share of total property crimes cleared by arrest abruptly falls.

The above mechanisms point to both immediate and gradual implications for lethal use-of-force, consistent with the baseline results. In the short-term, the protests reduce the number of opportunities to use lethal force by decreasing the quantity of policing. In the long-term, protests associate with a gradual rollout of institutional reforms. Such a

---

<sup>1</sup>Because the data are stacked by cohort, the estimates are robust bias arising from treatment effect heterogeneity over time or place (Baker et al., 2021).

<sup>2</sup>I gauge robustness to population screens up to 250,000 and choice of dataset.

<sup>3</sup>I also assess the influence of time-variant controls.

<sup>4</sup>As an alternative, I demonstrate that population decile-time fixed effects capture these pretend differences, suggesting the culprit may indeed be population-driven, time-variant measurement error. Furthermore, the results are robust to using census place-specific linear time trends and using only post-2013 data from Mapping Police Violence (MPV), which is tenable to changes in under-reporting.

distinction, however, is made less relevant to this paper by the continuous administration of protests during the post-treatment period.

Viral videos and protests are often concurrent; therefore, the baseline estimates may misattribute the impact of viral videos to protests. I provide falsification tests by leveraging the timing and location of both videos and protests; if videos reduce lethal use-of-force in places without protests, then I am likely overstating the impact of protests; likewise, if protests do not reduce lethal use-of-force without video, then I am less confident in any direct protest effect. These falsification tests are passed when using the entire sample and when pooling only the 30 most prominent police scandals from 2014-2019.<sup>5</sup>

The baseline estimates may use invalid control places even after reweighing due to noise in the lethal use-of-force measure and do not account for spillover effects onto nearby jurisdictions. I address both concerns by aggregating from census places to counties. This aggregation alleviates concerns regarding local spillover effects by extending treatment status to nearby jurisdictions.<sup>6</sup> This aggregation also allows for contiguous county estimates, which accounts for spatial heterogeneity. The principal result hold. There is also a nonnegligible increase in the immediate effect of the protests. Protests, therefore, may impact the behavior of both local and nearby police officers, attenuating shorter-run estimates, which is consistent with research on the spillover effects of police scandals (e.g. [Cheng and Long, 2018](#); [Premkumar, 2020](#)). Finally, I show that the synthetic unit weights tend to emphasize contiguous counties, albeit fewer of them, which bolsters confidence in the synthetic DID estimator.

This article contributes to the small but proliferating literature on BLM by providing the most comprehensive interrogation of BLM's impact on lethal use-of-force so far, the animating purpose of the movement ([Mazumder, 2019](#); [Sawyer and Gampa, 2018](#); [Trump et al., 2018](#); [Skoy, 2020](#)). Research has found that BLM has reduced white racial prejudice ([Mazumder, 2019](#); [Sawyer and Gampa, 2018](#)). This finding potentially links protests with lethal use-of-force, as neighborhood racial bias is associated with disproportionate lethal use-of-force ([Hehman et al., 2018](#)). However, the direct relationship between BLM and lethal force has not been adequately studied. [Skoy \(2020\)](#) is the first and only study thus far to investigate BLM's effect on lethal use-of-force. Using a monthly panel of states, the author finds that BLM protests reduce fatal police interactions in the proceeding month but does not explain why. The study's empirical methodology is problematic for aggregating above the level of treatment, as BLM's brunt will likely be felt at a city level, influencing local policy,

---

<sup>5</sup>Demonstrations preceded the viral videos in over half of the 30 case studies. In many of these cases, media reports suggest the videos were made public due to public pressure from the protests. It thus seems reasonable to infer that the effect of the video is part of the causal channel of protests for these specific cases.

<sup>6</sup>If spillover effects were to have occurred nationally, then their effect would be absorbed by the time fixed effects and, thus, would not be a concern.

the local police agency, and residents' attitudes. Worse, the omission of local differences renders their identifying assumption dubious.

My article also contributes to the literature on the effect of scrutiny or public monitoring on policing behavior (Premkumar, 2020; Devi and Fryer Jr, 2020; Ba and Rivera, 2019; Kochel, 2019; Long, 2019; Shjarback et al., 2017). The most recent research in this literature has focus on community trust (Kochel, 2019) and the Ferguson Effect. An emerging literature has found evidence consistent with such an effect (Devi and Fryer Jr, 2020; Premkumar, 2020; Cheng and Long, 2018); however, one did not (Pyrooz et al., 2016). This article not only provides one of the largest tests for depolicing in the literature but also offers compelling evidence that protests, not videos in-of-themselves, are a driving force.

This article fills a gap in the literature on how racial justice protests shape police behavior. Cunningham and Gillezeau (2018) finds the African American uprisings during the 1960s, often a response to police violence, resulted in an increase in police homicides against nonwhite residents. While taking place half a century earlier, this dismal result highlights how the *a priori* sign of BLM's effect on police homicides is ambiguous. Unlike the 1960s when police responded to protests with more aggressive policing, this paper suggests police have reacted to BLM with less policing in general.

Finally, this paper broadly contributes to the empirical literature on the effectiveness of protesting by showing how recent innovations in the DID literature can inoculate against endogeneity endemic to the research question when common instrumental variables are invalid (Mazumder, 2019; Cunningham and Gillezeau, 2018; Van den Broek et al., 2017; Madestam et al., 2013).

The rest of the article is structured as follows. Section II explains the empirical methodology and data sources, detailing how each estimator mitigates one or more of the four hurdles listed above. Section III presents the main empirical findings on BLM impacts on lethal use-of-force. Section IV gives estimates for how the protests alter local police agencies and public scrutiny of police, which are likely mechanisms behind the primary finding. Section V concludes. Finally, the location of publicly available replication files is given in Section VI.

## II. Methodology and Data

There are several potential challenges in estimating the effect of BLM protests on police homicides. First, there is currently no federal database with credible data on police killings, a long-standing problem (Fyfe, 2002; Banks et al., 2015; Klinger et al., 2016; White, 2016). The absence of data has been filled by nonprofit and media organizations who likely undercount the

true number of police homicides when incidents go unreported. Time-variant improvements in reporting compound this issue. In 2013, Fatal Encounters initiated its reporting system and retroactively recovered older incidents. This change implies undercounting may become more severe as years retrocede 2013. The gains in reporting may also have been more substantial for large cities, as online records are more available in larger cities than smaller cities. Since large cities are also more likely to have BLM protests ([Trump et al., 2018](#)), reporting system improvements may violate the parallel trends assumption of DID estimators without general differences in police killings.

Second, a city's characteristics may influence the likelihood of both BLM protests and lethal force, some of which may be unobservable. For example, [Trump et al. \(2018\)](#) associated BLM protests with poverty, educational attainment, population size, police killings, the democratic vote share, the black population share, and past incidents of lethal use-of-force by the police. Even if not directly related to police homicides, unobserved correlates of protests could induce omitted variable bias if correlated with determinants of lethal force.

Third, since BLM protests are motivated by police killings, police homicides likely rise before BLM protests, but only in cities where the protests occur ([Trump et al., 2018; Skoy, 2020](#)). This pre-trend difference would break the identifying assumption of DID estimators used in some related research ([Mazumder, 2019; Cunningham and Gillezeau, 2018](#)).

Forth, the prevalent rainfall instrumental variable design, which uses rainfall to isolate exogenous variation in protest participation, is not valid. While the relevance of rainfall for protest participation is well established ([Zhang, 2016](#)), the exclusion restriction is unlikely *a priori*. Rainfall likely determines lethal force regardless of protest participation for the same reason rain alters protest turnout; people go outside less when it's raining. The weather may also affect lethal use-of-force indirectly, as rising temperatures and low rainfall incite aggression and violent crime ([Carleton and Hsiang, 2016](#)).

I address these problems with a stacked DID design that leverages variation in BLM protests' location and timing to uncover the BLM's effect. I contrast four different estimators to address the above issues.

## II.A. Empirical model

The primary aim of this study is to evaluate the impact of BLM protests on incidents of lethal use-of-force. The baseline model is a stacked DID design with two-way fixed effects:

$$\frac{Y_{c,i,t}}{N_{c,i,t}} = \mu + \sum_{k=-4}^4 \beta_k D_{k,c,i,t} + X'_{c,i,t} \kappa + \alpha_{c,i} + \delta_{c,t} + \epsilon_{c,i,t} \quad (1)$$

where  $Y$  is the count of lethal use-of-force and  $N$  is the normalization variable (none, population, officers, violent crime, or total arrests) in census place  $i$  during time  $t$  (quarterly) within cohort  $c$ ,  $D_k$  takes value one during event-year  $k$  for places that have protests and zero otherwise, and  $X$  is a vector of time-variant controls. The dataset is a ‘stack’ of cohorts. To be clear, each cohort includes all treated places that witness their first BLM protest during the same quarter and all control places. Control places are any cities that do not have a BLM protest during the entire sample. Meaning, treated places will be included in only one cohort, but control places will be included in every cohort. The time variable for each cohort is centered at the quarter of the first protest for both treated and control units. So  $t = 0$  always corresponds to the first protest for treated cities. The benchmark specification controls for population flexibly by fitting a linear control for the population for each cohort-population decile. Stacking by cohort requires cohort-place fixed effects  $\alpha_{c,i}$  and cohort-time (quarterly) fixed effects  $\delta_{c,t}$ . The standard errors are clustered by census place since this is the level protests are assigned. The standard errors account for possible correlation within a city in the changes in lethal use-of-force.

This DID model identifies the effect of the BLM protest on lethal use-of-force if police homicides would move in parallel between places with and without protests had the protests never occurred. This assumption holds if all determinants of BLM protests are either observed, time-invariant, or common across all places:  $E(\epsilon_{c,i,t} | \{D_k\}_{k=-4}^4, X_{c,i,t}, \alpha_{c,i}, \delta_{c,t}) = 0$ . While not directly testable, I use a common practice of assessing the parallel trends assumption with leading terms. Specifically, this specification allows for trends to deviate four years before a protest occurring ( $\beta_{-4}, \beta_{-3}, \beta_{-2}, \beta_{-1}$ ); detecting a difference during these years would indicate a violation of the parallel trends assumption.

Equation 1’s slightly non-standard way of delineating event time implies that  $\beta_k$  denotes the relative difference of year  $k$  to 5 or more years prior to treatment (the omitted category), not the year leading up treatment (the standard omitted category). Indeed, this nonstandard base category requires a second step for interpretation. I estimate the percentage change in police homicides per normalizing variable by dividing the estimated  $\beta_k$  from Equation 1 by the average lethal use-of-force per normalizing variable among places exposed to BLM protests one year prior to the first protest ( $\bar{b}_{-1}$ ). The annual percentage change in police homicides in year  $k$  is  $\frac{\beta_k - \sum_{k=-4}^{-1} \beta_k / 4}{\bar{b}_{-1}}$  and the average, annual percentage change is:

$$\% \Delta \text{Lethal Force} = \frac{\sum_{k=0}^4 \beta_k / 5 - \sum_{k=-4}^{-1} \beta_k / 4}{\bar{b}_{-1}}.$$

By subtracting the average of pretreatment coefficients, the estimates now have the same

interpretation as the standard method but with a level difference.<sup>7</sup> This two-step procedure has advantages over the traditional method. First,  $\beta_{-1}$  can be used as a placebo test since it is not fixed to zero. Second, the results are less sensitive because the estimates are centered over the entire preintervention event-window rather than over one year. Third, retaining preintervention data prior to the event-window bolsters credibility when using unit-specific time trends, as they require adequate pre-intervention data to capture pre-existing trends, a common robustness test.

Another interesting statistic is the total change in lethal use-of-force attributable to BLM protests, which is the product of the average quarterly change in lethal use-of-force after demonstrations, the total number of time-places exposed to at least one protest ( $e$ ), and the average normalizing variable among places exposed to BLM protests one year before the first protest ( $\bar{n}_{-1}$ ):

$$\Delta \text{Total Lethal Force} = \% \Delta \text{Lethal Force} * \overline{\sqrt{N}}_{-1} * e * \bar{N}_{-1}.$$

This estimator has a major advantage over the more standard staggered DID estimator. Even when the parallel trends assumption holds, the staggered DID (or event study) can be biased from heterogeneity (Goodman-Bacon, 2018; Sun and Abraham, 2020; Callaway and Sant'Anna, 2020). In particular, if there is heterogeneity over time, the staggered DID uses already-treated units as control units, creating bias from negative weighting. This problem is resolved by stacking because, when treatment timing is aligned by cohort, already treated units cannot be selected as controls (Baker et al., 2021; Goodman-Bacon, 2018).

To gauge the robustness of the results, Equation 1 is approximated with four different estimators, detailed below, with the following form:

$$(\hat{\mu}, \hat{\beta}, \hat{\alpha}, \hat{\delta}) = \arg \min_{\mu, \beta, \alpha, \delta} \sum_i \sum_t \left( \frac{Y_{c,i,t}}{N_{c,i,t}} - \mu - \sum_{k=-4}^4 \beta_k D_{k,c,i,t} - \alpha_{c,i} - \delta_{c,t} \right)^2 w_{c,i,t} \quad (2)$$

If Equation 1 is correctly specified, then all four estimators are consistent. If Equation 1 is incorrectly specified, then some of the alternative estimators are still consistent given the weights are appropriately penalized. Thus, similarity between the estimators is consistent with Equation 1 being the correct specification, which would bolster confidence in the results.

---

<sup>7</sup>For balanced event studies with two-way fixed effects in general, if the annual change in time  $k$  is estimated as  $\beta_k - \beta_{-1}$ , then it is equivalent to the standard saturated model with data in the prior omitted category dropped (pre-event window).

### ***II.A.1. Ordinary Least Squares***

The benchmark estimator is ordinary least squares (OLS) without normalization of police homicides.

$$w_{c,i,t} = 1 \quad \text{and} \quad N_{c,i,t} = 1 \quad (3)$$

The estimator is identified if all BLM protest determinants are either time-invariant or common across all places.

### ***II.A.2. Per Capita Population Weighted Least Squares***

The second estimator is per capita population weighted least squares (WLS), which accounts for population-driven heteroscedasticity.

$$w_{c,i,t} = \sqrt{\text{Population}_{c,i,t}} \quad \text{and} \quad N_{c,i,t} = \text{Population}_{c,i,t} \quad (4)$$

Like OLS, the estimator is consistent if BLM protests determinants are time-invariant or common across all places, thus its contrast to Equation 3 gives a diagnostic test for model specification (Solon et al., 2015). The key difference is that weighting by population places more weight on observations with greater precision if the media neglects protests in less populated areas (Madestam et al., 2013).

### ***II.A.3. Doubly Robust Inverse Probability Weighting***

The previous specifications assume that all characteristics associated with BLM protests either do not change over time or the change is standard across places. Therefore, unobserved correlates of both protests and lethal force could contaminate such estimates. I inoculate against omitted variable bias with lasso regularized inverse-probability weighting using pre-protest controls. The controls are established correlates of either BLM protests or police homicides;<sup>8</sup> they include the following: local police agency characteristics and policy such as officer race, officer gender, full-time officer employment, race and gender of supervisors,

---

<sup>8</sup>A large body of work has investigated lethal use-of-force determinants. Police department policies are related to use-of-force (Terrill and Paoline, 2017), such as use-of-force reporting requirements (McElvain and Kposowa, 2008) and body-worn cameras (Ariel et al., 2015). Officer race, gender, education, and experience have been linked with use-of-force, and black people are more likely to be subjected to non-lethal force than white people (Ba et al., 2021; Ridgeway, 2020; Fryer, 2019; Tregle et al., 2019; Paoline III and Terrill, 2007; Hoekstra and Sloan, 2020; Ross, 2015). However, there is still debate over a racial disparity in lethal force because the existence depends on normalization (Menifield et al., 2019; Tregle et al., 2019; Buehler, 2017; Cesario et al., 2019; Tregle et al., 2019; Fryer, 2019). Racially biased communities tend to have higher rates of police homicides (Hehman et al., 2018), as with settings with mid-level violent crime rate (Klinger et al., 2016). Areas with a high proportion of black-on-white homicides experience a higher rate of police homicides, especially by white police officers (Legewie and Fagan, 2016). Some research shows the importance of police training (Joshua et al., 2007; Donner and Popovich, 2018).

turnover, community policing, authorized use-of-force equipment, use-of-force documentation requirements, and unionization; crime such as property crimes, violent crimes, homicides and assaults and murders of officers; demographics such as race, educational attainment, poverty, labor force participation, and unemployment; geographic controls such as population density and city size; democratic vote share in the 2008 presidential election; historic civil rights protests and hate crimes; consent decrees.

The procedure has five steps. First, the control variables are collapsed into 2013 means by census place.<sup>9</sup> Second, ten datasets are imputed using the chained ten nearest neighbor mean. Third, lasso logistic regression with the so-called plugin penalty is used to estimate the propensity score using the stacked imputed datasets:

$$D = X_{2013}\theta + Y_{Pre}\gamma + v$$

where  $D$  is the treatment indicator,  $X$  is a matrix of the control variables listed above, and  $Y$  is a matrix of pretreatment annual pretreatment outcome means. Fourth, the propensity scores are averaged over the imputed datasets weighted by the fraction of missing data.<sup>10</sup> Fifth, inverse probability weights are constructed using the propensity scores  $\hat{P}$ :

$$w_{c,i,t} = \hat{\omega}_i = 1\{D_i = 1\} \left( \frac{\hat{\mathbb{E}}(D_i)}{1 - \hat{\mathbb{E}}(D_i)} \right) + 1\{D_i = 0\} \left( \frac{\hat{P}_i}{1 - \hat{P}_i} \right) \quad \text{and} \quad N_{c,i,t} = 1. \quad (5)$$

This estimator is doubly robust. The estimator is consistent if the logit model for BLM protests is correctly specified and Equation 1 is misspecified. Alternatively, the estimator is consistent but inefficient if Equation 1 is correctly specified, and the logit model is misspecified (See [Imbens and Wooldridge, 2009](#)).

#### ***II.A.4. Synthetic Difference-in-Differences***

The greatest threats to identification are violations of the parallel trends assumption from protest responding to prior killings or the 2013 police homicide reporting improvements in large cities relative to smaller cities. I reduce the potential for such bias by adapting [Arkhangelsky et al. \(2019\)](#)'s synthetic DID estimator.

The main advantage of synthetic DID is double-bias reduction. Unlike the synthetic control method, which matches the pretreatment outcomes over units with time fixed effects, the synthetic DID approach also balances control outcomes over time and adds unit fixed

---

<sup>9</sup>Since 2013 is the last year before the first cohort of BLM protests, this ensures the control variables do not contaminate the impact of the protests.

<sup>10</sup>For a discussion of stacked multiple imputation for regularized regression, which is the imputation procedure used here, see [Wan et al. \(2015\)](#) and [Zhao and Long \(2017\)](#).

effects. Put simply, if the unit-weights do not fully balance the underlying signal in the pretreatment period, time-weights may balance the remainder. A secondary benefit is avoiding the need for testing pre-trends, which may bias published research (Roth, 2019). Like the IPW estimator, synthetic DID is doubly robust.

This procedure has three steps and is carried out for each cohort separately. First, IPW weights are estimated that match the pretreatment outcomes between places with and without protests,  $\omega_i$ . The cohort-place propensity scores are the ridge-penalized predicted values of the equation:

$$D = Y_0\eta + v$$

where  $Y_0$  is a matrix of pre-protest lethal use-of-force demeaned by cohort-place and cohort-time. The ridge penalty is selected with 10-fold cross-validation. Using the cohort-place propensity scores  $\hat{P}_i$ , the cohort-place IPW weights are:

$$\hat{\omega}_{c,i} = 1\{D_{c,i} = 1\} \left( \frac{\hat{\mathbb{E}}(D_{c,i})}{1 - \hat{\mathbb{E}}(D_{c,i})} \right) + 1\{D_{c,i} = 0\} \left( \frac{\hat{P}_{c,i}}{1 - \hat{P}_{c,i}} \right).$$

The unit weights are next scaled to sum to the number of treated units in the cohort. Second, IPW weights are estimated that match the pre- and post-protest lethal use-of-force of the control cohort-places (demeaned by cohort-place and cohort-time). The time propensity scores are the ridge-penalized predicted values of the equation:

$$D^* = Y_{\text{Control}}\eta + \nu$$

and are denoted as  $\hat{P}^*$ .  $D^*$  is an indicator for  $t \geq 0$ . The ridge penalty is again selected with 10-fold cross-validation. The cohort-time IPW weights are:

$$\hat{\lambda}_{c,t} = 1\{D_{c,t}^* = 1\} \left( \frac{\hat{\mathbb{E}}(D_{c,t}^*)}{1 - \hat{\mathbb{E}}(D_{c,t}^*)} \right) + 1\{D_{c,t}^* = 0\} \left( \frac{\hat{P}_{c,t}^*}{1 - \hat{P}_{c,t}^*} \right)$$

The time weights are next scaled to sum to the number of post treatment periods in the cohort. Third, the final weights are the product of the cohort-place and cohort-time weights:

$$w_{c,i,t} = \hat{\omega}_{c,i}\hat{\lambda}_{c,t} \quad \text{and} \quad N_{c,i,t} = 1. \tag{6}$$

Identification now assumes either selection is on fixed effects or the weights are correctly penalized.

## II.B. Data and Sample Construction

### II.B.1. Lethal Force Data

There is currently no federal database with reliable police killings data (Fyfe, 2002; Banks et al., 2015; Klinger et al., 2016; White, 2016). Nonprofit and media organizations have filled the absence of data. Through a combination of crowdsourcing, freedom of information act requests, and media coverage, public datasets are now available on police homicides including [KilledByPolice.net](#), [The Homicide Record](#) by the Los Angeles Times, [Mapping Police Violence](#) (MPV), the [Washington Post](#), the [Counted](#) by the Guardian, and [Fatal Encounters Dot Org](#).

Legewie and Fagan (2016) analyse the quality of the latter three sources, which are widely used (e.g. Trump et al., 2018; Skoy, 2020; Cesario et al., 2019; Nix et al., 2017).<sup>11</sup> Of the 1147 total police killings in 2015, the authors find Fatal encounters were missing 33 incidents, the Guardian was missing 49 incidents, and the Washington Post was missing 184 incidents. While Fatal Encounters was the most complete, the information on race was subpar.

Police homicides are measured as fatal encounters with police resulting from asphyxiation, bludgeoning, a gunshot, pepper spray, or a taser that are not suicides. The benchmark estimates use D. Brian Burghart's nonprofit [Fatal Encounters Dot Org](#). The organization operates three main methods for collecting data: 1) Paid researchers (85% of data), 2) Public records requests, and 3) Crowdsourcing. Paid researchers aggregate data from other sources listed above. All data are then verified by a principal investigator, cited, and checked against published sources. The dataset is updated regularly and begins in 2000. In 2013, Fatal Encounters initiated its reporting system and retroactively recovered older incidents. This change implies that undercounting may become more severe as years retrocede 2013. The Fatal Encounters data are detailed. For each police-involved fatality, they describe the incident, the address of the death, but the information on race, weapons, and disposition of death are worse than MPV.

MPV is the highest quality data on lethal force from 2013 to 2019. The organization gathers data from the other previously mentioned databases, improving their quality and completeness by “searching social media, obituaries, criminal records databases, police reports, and other sources to identify the race of 91 percent of all victims in the database.” MPV also has detailed information on the alleged arming of the victim. However, MPV does not have data before 2013, implying pre-trend differences in police homicides before BLM protests cannot be tested, a significant drawback. Hence, MPV is only used to see if the estimates hold using alternative data with less measurement error or vary by race or alleged arming of the victim.

---

<sup>11</sup>See [Bor et al. \(2018\)](#) and the ensuing correspondence for a discussion of the quality of MPV.

Fatal Encounters' definition of police homicides is too broad: all lethal interactions with police, whether on- or off-duty, including suicides. However, MPV only includes cases where "a person dies as a result of being shot, beaten, restrained, intentionally hit by a police vehicle, pepper-sprayed, tasered, or otherwise harmed by police officers, whether on-duty or off-duty." Appendix A Figure A.1 displays the proportion of total fatal encounters by cause of death during 2013-2019 and contrasts Fatal Encounters with MPV. Gun force is particularly lethal; however, the injury may have been self-inflicted, thus not lethal use-of-force. The Fatal Encounters data suggests gunshots account for 69.13% of the 10725 deaths. However, according to MPV, there were only 7642 police homicides, 95.21% of which resulted from a gunshot. Vehicle-related deaths make up most of the discrepancy. Police homicides are restricted to fatal encounters from asphyxiation, bludgeoning, a gunshot, pepper spray, or a taser that are not suicides to mitigate the discrepancy.<sup>12</sup> The restricted definition misses 39 cases in the MPV data but makes the case-of-death distribution and total cases similar, bolstering confidence in the measure's quality.

### ***II.B.2. Black Lives Matter Protest Data***

The BLM protests data builds from [published data](#) by [Trump et al. \(2018\)](#) who use a rolling web-search to count the number of protests by census place from August 9<sup>th</sup>, 2014 through August 9<sup>th</sup>, 2015. I then web-scrape data from August 10<sup>th</sup>, 2015 through 2019 from a [website](#) maintained by Alisa Robinson. She is a graduate of the political science department at the University of Chicago. Her data is publicly available through a Creative Commons license.

This paper is concerned with the effect of BLM protests against police violence, particularly the public gathering of individuals. For this reason, I exclude online demonstrations, protests by professional athletes, protests against presidential candidates, or protests against conservative talks at universities.

### ***II.B.3. Data for Control Variables***

The decennial census is used for the census place geographic size and the number of houses, while the annual intercensal census is used for population ([U.S. Census Bureau, 2019](#)). I use the 2013 five-year [American Community Survey](#) for data on poverty rate, labor force participation rate, unemployment rate, full-time employment rate, the black population share,

---

<sup>12</sup>A priori, the inclusion of deaths that do not result from lethal force may attenuate the estimates, which is why the baseline estimates omit vehicle related deaths, the main force behind the MPV and Fatal Encounters discrepancy. That said, some police departments during this sample initiated no-chase policies to reduce vehicle-related deaths. For this reason and others, vehicle-related deaths are no less important from a welfare perspective. Thus, I report estimates using the MPV data that do not omit lethal force incidents by cause of death in Appendix A Table A.1. The estimates do not meaningfully change.

the black poverty rate, and educational attainment measures, including the portion of people with less than a high school, some college, or college education ([U.S. Census Bureau, 2013](#)). The year 2013 is chosen to ensure control variables do not contaminate the effect of the BLM protests.

I next use Jacob Kaplan's concatenated files of the [Uniform Crime Reporting](#) data to obtain data on property crime, violent crime, murder, assaults on police, and felonious police deaths ([Kaplan, 2020](#)).

To measure a city's protest history before BLM, I use the [Dynamics of Collective Action](#) dataset. The publically available dataset counts the number of protests and hate crimes in each city based on media reports. The dataset codes each demonstration according to the participants and demand of the action. The data counts the number of pro-black civil rights protests, pro-anti-police brutality protests, black initiated protests, and racist events, including hate crimes and protests against the civil rights of racial minorities.

The data for police agency characteristics comes from multiple sources. The [Annual Survey of Public Employment and Payroll](#) (ASPEP) is used for measures of the annual number of police officers and the average wage for a police officer for each place. I use the [Law Enforcement Management and Administrative Statistics](#) (LEMAS) 2013, 2016, and 2016 Police Body-Worn Camera Supplement to complement the ASPEP data on police wages and the number of police; the average value of the two sources is taken for each place. The LEMAS also provides information on agency characteristics, including officer demographics, unionization, use-of-force reporting, authorized equipment, police training, body-worn cameras, and community policing initiatives ([DOJ, 2015](#); [DOJ, 2020](#); [DOJ, 2019](#)). Police agency characteristics are linked to census places using the [Law Enforcement Agency Identifiers Crosswalk, United States, 2012](#) ([BJS, 2018](#)).

Last, the city level [democratic vote share](#) in the 2008 presidential election is taken from [Einstein and Kogan \(2016\)](#).

#### ***II.B.4. Sample and Covariate Balance***

The final dataset includes any census place with a population of at least 20,000. I collapse the data into quarterly counts of BLM protests and police homicides for each census place from 2000q1 until 2019q4.

Appendix A Figures A.2 and A.3 display the dispersion of protests. Initial protests concentrate in urban areas during the wake of the deaths of Mike Brown and Eric Garner during the summer of 2014 and Alton Sterling and Philando Castile during the summer of 2016. This concentration implies that the variation in treatment timing is limited, limiting the prospects for identification. The first cohort includes 31.6% of events (2014q3), the second

includes 22.21% (2014q4), the ninth includes 12.25% (2016q3), the tenth includes 5.16% (2016q4), and the remaining 28.38% of events are more-or-less uniformly distributed over the years 2015 and 2017-2019; the particular quarters show a seasonal trend. The protests are persistent. If a rally occurs, then, on average, seven more occur over the subsequent five years. Three demonstrations during the first year of protest are usual, with one or two following demonstrations annually (see Appendix A Figure A.4). Therefore, it is important to be cognizant of the continuous administration of protests over the entire course of the post-treatment period. The impacts of the protests hence may be dynamic, accumulating with time.

Table 1 reports the covariate balance between places with at least one BLM protest (treated group) and places without a BLM protest (control group) under each weighting scheme. Columns 1 and 2 show the unweighted means for each control variable by treated status. The results are consistent with the findings of Trump et al. (2018); places with at least one BLM protest tend to have a higher poverty rate, a larger black population share, a higher black poverty rate, more college education, and a larger population. These differences improve for all weighting schemes, and the covariate balance is strongest with IPW control weights; however, population remains exceptionally unbalanced, as does the black population share and black poverty rate. Furthermore, the joint F-test suggests the treated and control groups are observably different for all weight schemes. This imbalance raises concerns regarding unobserved differential trends between treatment and control places; put plainly, the experimental design compares larger cities (mean population of 243,000) to smaller cities (mean population of 107,000).

To mitigate the problem of population imbalance, all specifications flexibly control for population with a linear cohort-population decile interaction. Furthermore, I assess the robustness of the primary finding to controlling for population decile-time interaction to control for population-driven heterogeneity.<sup>13</sup> This test ensures comparison is drawn between places with and without protests within the same population-decile. The final estimate is an average across population groups. I also report results that incrementally increase the population screen from 20,000 to 250,000. The primary finding does not meaningfully change in any of these cases.

---

<sup>13</sup>I also test sensitivity to the inclusion of cohort-place linear time trends and other time-variant controls.

### III. Results

The validity of a DID design rests on a parallel trend assumption. Given the observable differences detailed in Section II.B.4, the trends in lethal use-of-force are remarkably parallel until 2013, which is around one year before BLM for most cohorts (see Figure 1). Between 2013 and 2014q3, however, treated cities, which are relatively large, racially diverse, college-educated, and democratic party leaning, experienced a decline in police homicides, whereas control cities did not. The two most likely explanations are measurement error in lethal use-of-force and confounding from pre-BLM protests. Indeed, the negative pretends in lethal force from 2013-2014 corresponds with both the improvement in lethal force reporting that was likely stronger in larger cities (2013 onward) and the racial justice protests sparked by the acquittal of George Zimmerman (summer 2013); both call the parallel trends assumption into question.

Following BLM protests, lethal use-of-force fell by 15.8% (s.e.=0.046) on average, conditional on population and fixed effects (see Table 2 Column 1). If the model is correct, then BLM protests were responsible for approximately 300 fewer people killed by the police from 2014 through 2019. The payoff for protesting is substantial; around every 5 of the 1,724 protests in the sample corresponds with approximately one less person killed by the police over the following years, depending on specification. The police killed around one less person for every twelve hundred participants.

The baseline estimate should be interpreted with caution. As with all DID applications, these estimates' validity rest on assuming parallel trends between places with and without exposure, had protests never occurred. Figure 2 gauges the validity of the assumption by allowing trends to deviate for four years preceding BLM. No pre-trend difference is detected; however, the test could be underpowered. For already stated reasons, the trend difference two years before BLM creates concern albeit statistically insignificant.

Protests and lethal force are more likely to be under-reported in lower population areas since both data sources rely on media reporting. Normalizing police homicides by population and weighting by population accounts for the population-driven variance created by this under-reporting. Intuitively, normalizing by population allows for easier comparisons in lethal force, and weighting by population emphasizes observations less prone to measurement error. Column 2 and Figure 2b report the population-weighted per capita regression. The estimates are slightly larger; BLM protests corresponded with a 17.7% (s.e.=0.052) reduction in lethal use-of-force. The negative pre-trend diminishes but remains.

Cities that experienced protests tended to be larger than cities without them. They also tended to have a greater share of Black, college-educated residents and individuals

who voted Democratic in the 2008 presidential election. To adjust for these differences, Column 3 and Figure 2c report the WLS regression estimate where weights balance the inverse probability of having at least one protest between cities based on their average 2013 characteristics. The estimates are again larger; BLM protests associated with a 21.6% (s.e.=0.065) reduction in police homicides. The adjustment for covariate imbalance does not alleviate the negative pre-trend; this may be because the adjustment fails to adequately balance observable characteristics, as described in Section II.B.4, making unobserved differences likely.

The most severe threat to identification is a violation of the parallel trends assumption; protests could reflect responses to past incidents of police homicides, lethal force measurement error correlated with population and time, or greater protest likelihood in cities with already high police scrutiny. To inoculate against violations of the parallel trends assumption, Column 6 and Figure 2d report the synthetic DID estimates that weight by the product of the cohort-place and cohort-time inverse probability weights. The cohort-place weights balance the number of police homicides over the four years before the cohort's first BLM protests, between places with and without an eventual protest. The cohort-time weights balance the data according to the signals in the control outcomes. The estimates do not meaningfully change; BLM protests associated with a 14.1% (s.e.=0.050) reduction in police homicides. This estimator eliminates the negative pre-trend.

The results stand in stark contrast to [Cunningham and Gillezeau \(2018\)](#). I provide evidence in Section IV.B that may explain why there was a reduction, rather than increase, in use-of-force following BLM protests. The police began responding to protests with less rather than more aggressive policing.

### III.A. Robustness checks

The baseline estimates use Fatal Encounters data (available from 2000-present) rather than higher-quality data from MPV (available from 2013-present) to test for pre-trend difference in lethal use-of-force. By doing so, the parallel trends assumption is susceptible to population-driven, time-variant measurement error from improvements in reporting in larger cities. The baseline estimates also omit police homicides with specific causes of death to safeguard against attenuation from including fatal police encounters unrelated to use-of-force. Table A.1 reports estimates using MPV data to address both issues. The principal results do not meaningfully change, providing much-needed assurance that the 2013 improvements in lethal force reporting are not behind the finding.

The MPV data allows for subsetting lethal force incidents by race and unarmed status. Race is an essential factor in use-of-force. Not only are black people more likely to be subjected to non-lethal force than white people ([Alpert and Macdonald, 2001](#); [Fryer, 2019](#);

Tregle et al., 2019), but white officers use gun force at a rate well above black officers when dispatched into black neighborhoods (Hoekstra and Sloan, 2020). The fall in lethal force following protests was greatest for the unarmed and was similar for black people and white people (see Table A.1).

So far, the specifications use an indicator for having at least one BLM protest to capture the impact of protesting, which may fail to capture the protests' intensity. Table A.2 reports the percentage change in lethal force from BLM protests sub-setting treated units by quartiles of maximum protests size and frequency; the set of control units is constant. Since major changes following minor events are unlikely, these estimates also provide a falsification test. The fall in lethal use-of-force was extensive and precise in the fourth quartile of maximum protest size or frequency. Places without a protest of this size did not significantly decrease police homicides for most specifications. However, the largest fall is not found in the fourth frequency quartile but the second, which is somewhat concerning (given the imprecision).

One may be concerned that the baseline estimates do not adequately account for spatial heterogeneity or the spillover effects of protests onto nearby jurisdictions. Appendix B addresses both concerns by aggregating from census places to counties. This aggregation allows for contiguous county estimates, which account for spatial heterogeneity. This aggregation also addresses local spillover effects by extending treatment status to nearby jurisdictions. If spillover effects were to have occurred nationally, then their effect would be absorbed by the time fixed effects and, thus, would not be a concern. The principal results hold at a county level using ordinary least squares, per capita population weighting, contiguous counties, or synthetic DID. There is also a nonnegligible increase in the immediate effect of the protests, suggesting local spillovers may attenuate the shorter-run impacts of protests on lethal use-of-force. This finding points to protests impacting the behavior of both local and nearby police officers, which is consistent with related research (e.g. Cheng and Long, 2018; Premkumar, 2020).

Protests did not meaningfully alter lethal force when normalized by crime (see Appendix A Table A.3), suggesting a decrease in encounters that could lead to police homicides drove the lethal force reduction rather than a change in use-of-force propensity. The benchmark results omit all census places with a population below 20,000. Appendix A Table A.4 shows robustness to population screens up to 250,000. Appendix A Table A.5 and Appendix A Figure A.5 show robustness to including time-variant controls, cohort-time fixed effects, cohort-place linear time trends, and weighting choice.

## IV. Why Does Lethal Force Fall After Black Lives Matter Protests?

### IV.A. Body-Worn Cameras and Other Police Agency Reforms

The results thus far indicate BLM protests reduced lethal use-of-force locally but do not explain why. One way BLM may impact lethal use-of-force is by pressuring local police agencies to change. I gauge these mechanisms with a DID design. I create a two-period panel of police agencies using all agencies that responded to both the 2013 and 2016 Law Enforcement Administrative Statistics (LEMAS) for all outcomes other than body-worn cameras. For body-worn cameras, I construct a quarterly panel of agencies stacked by cohort analogous to Equation 1 using the 2016 Body-Worn Camera supplement, which elicits the month each agency acquired body-worn cameras. Both panels are linked to census place level protests from 2014-2016 using the Law Enforcement Agency Identifiers Crosswalk, 2012. To be clear, one protest may be linked to multiple agencies.

BLM protests may influence lethal use-of-force by pressuring police departments to adopt body cameras, a likely use-of-force deterrent. While randomized control trials measuring the effectiveness of body cameras on use-of-force have had mixed results due to differences in pre-intervention use-of-force, policy/implementation, and research methods (White and Malm, 2020; Yokum et al., 2019; Peterson et al., 2018; Ariel et al., 2015; Jennings et al., 2015; Braga et al., 2018), a multisite randomized control trial found body-worn cameras reduce use-of-force when officers have limited discretion in turning them off (Ariel et al., 2016). A recent study by Kim (2019) forgoes the interval validity of a randomized control trial for the external validity of a DID estimator using the 2016 Body-Worn Camera supplement; the same dataset used in this article. The author finds a substantial decrease in use-of-force after agencies obtain body cameras.

Local police agencies may have responded to BLM protests by adopting body cameras to curtail force. Figure 3 shows that before protests, agencies that eventually had demonstrations were on a similar trend to agencies that did not have protests from 2014 to 2016. However, over the three years following protests, the share of agencies with body-worn cameras grew 68.3% (s.e.=0.119) for agencies with protest relative to agencies without protests, a strong indication that protests increased the adoption of body-worn cameras. Some qualitative evidence also supports this interpretation. Figure 4 shows that agencies with protests were significantly more likely than agencies without protests to report obtaining body-worn cameras to reduce use-of-force or to improve community perception.

If protests affect body camera policy, they should also impact the number of police

homicides with video recording. Over the five years following protests, there was a 231% (s.e.=0.410) increase in the number of police homicides with video recordings relative to places without protests, as shown by Appendix A Figure A.6. However, severe under-reporting of videos before 2015 may contaminate the estimates.<sup>14</sup> The relative number of agencies with body cameras during this time mitigates this problem; only ~20% of the sample police agencies had obtained body-worn cameras before 2015, yet ~50% had obtained body cameras by 2016.

A generous interpretation suggests that agencies with protests became more likely to obtain body-worn cameras and did so to reduce use-of-force and to improve community perception (not to improve evidence quality or reduce agency liability). [White and Malm \(2020\)](#)'s thorough review of body-worn camera randomized control trials shows that the mixed evidence is strongest when agencies obtain body-worn cameras because of a police scandal; scandal-ridden agencies have higher levels of unnecessary force, giving room for body-worn cameras to alter behavior. Thus, it is not surprising that agencies already operating with restrictive policies do not experience a reduction in use-of-force from body-worn cameras. It may be that body-worn cameras were relatively more effective in reducing lethal force when implemented as a response to protests of police scandals than when implemented for other reasons.

DID estimates suggest BLM protests' influenced the aspects of the local police agencies, see Table 3. Some police agencies assign officers to regular geographic patrols or encourage SARA-type (scanning, analysis, response, assessment) problem solving as a community policing initiative. BLM protests increased the number of officers with regular geographic patrols by 20.6% (s.e.=0.136), around forty officers (see Column 2). Column 3 reports a 57.5% (0.326) increase in the number of officers encouraged to engage in SARA-type problem-solving project. These results are consistent with police agencies expanding community policing due to pressure from BLM protests; however, the imprecision warrants a cautious interpretation. Column 4 reports a decrease in the number of black police officers by 6.7% (s.e.=0.033) and Column 5 shows a negligible impact on white officers. Because a reduction in black officers could correspond to a rise in use-of-force, especially in predominantly black cities (e.g. [Hoekstra and Sloan, 2020](#)), this is not consistent with the fall in lethal force. Column 6 reports a fall in officer experience (measured as the total number of offices less recruits), corresponding with more expected force ([Paoline III and Terrill, 2007](#)).

---

<sup>14</sup>In 2015, the Washington Post began documenting police shootings that were filmed by body cameras.

## IV.B. Public Scrutiny of the Police and Depolicing

If public scrutiny of the police damages police morale or community trust, a reduction in lethal use-of-force may result from a reduction in law-enforcement effort (the so-called Ferguson Effect).<sup>15</sup> Specifically, public scrutiny of the police may foster depolicing if officers experience disutility when interacting with critical members of the community and protesters. Officers may also perceive a higher cost for misconduct (Premkumar, 2020). Similarly, the protests may damage residents' trust in police (Kochel, 2019), increasing the number of crimes that go unreported. Both scenarios would lead to depolicing, potentially explaining why the police are less likely to use fatal force.

I test for the depolicing mechanism using three tests similar to Premkumar (2020) and Devi and Fryer Jr (2020) with the same regression specification as Equation 1 but with alternative outcomes. The quantity of policing will fall if public scandals reduce the community's willingness to report crimes and the marginal benefit of low-level arrests. I hence first test for a reduction in reported property crimes following protests, which are more susceptible to changes in community reporting and police effort than violent crimes. Second, if the share of property crimes cleared by arrest does not fall (does fall), then diminished police effort is unlikely (likely) because crime reporting is accounted for. Third, a decrease in the quantity of policing will lead to an increase in criminal offending. Therefore, a rise in criminal offending following the protests is consistent with depolicing. However, one cannot rule out a direct response to the protests. Accurate measurement of criminal offenses is difficult because low-level crime measures are also susceptible to police effort and community reporting. I hence follow Devi and Fryer Jr (2020) and use homicides to proxy for criminal offending, which inoculates against changes in police effort and community reporting.

Reported property crimes declined for all five years after protests (see Figure 5).<sup>16</sup> This finding is consistent with a persistent reduction in community trust or police effort. To disentangle these two mechanisms, the figure also depicts estimates for the share of reported property crimes cleared by arrest. In stark contrast to reported property crimes, the share of reported property crimes cleared by arrest abruptly fell after protests and then gradually reverted to its pre-trend. Taken together, these two results suggest the Ferguson Effect is temporary and abrupt, while the reduction in community cooperation is persistent and growing. The increase in homicides also bolsters the case for the public scrutiny mechanism; however, the one-year delay in the increase makes the result suspect. These results are

---

<sup>15</sup>Strict policies for body-worn cameras could theoretically cause depolicing (Wallace et al., 2018; Ariel et al., 2018). Therefore, protests could also create depolicing by influencing body-worn camera policy. In Section IV.A, I supplied evidence of agencies with BLM protests becoming more likely to obtain body-worn cameras.

<sup>16</sup>The results are also reported in Appendix A Table A.6, including additional crime measures.

consistent with the burgeoning literature on how public scrutiny of the police shapes crime, police effort, and community trust (Premkumar, 2020; Devi and Fryer Jr, 2020; Ba and Rivera, 2019; Kochel, 2019; Long, 2019; Shjarback et al., 2017).

#### IV.C. Chicken or the Egg: Are Protests Creating or Responding to Viral Videos?

These results raise a concern. If public scrutiny of the police drives the reduction in lethal force, am I overstating the role of protests? After all, the public scrutiny surrounding police scandals derives from both viral videos and protests. Indeed, the DID estimates may capture the impact of the scrutiny brought on by viral police scandals, not the effect of the protests themselves. Put differently, is protesting a response to viral police scandals, or is protesting responsible for the scandals going viral?

Evidence suggests a circular relationship between past and future video recordings. As documented in Appendix A, Figure A.6, a statistically significant increase in video recordings took place about one year before protests, followed by a stark increase in the number of video recordings over the five years after the first protests.

I document that protests preceded viral videos in over half of the 30 most prominent police scandals from 2014-2019 (see Appendix C). In many of these cases, media reports suggest the videos were made public due to public pressure from protests. It thus seems reasonable to infer that the effect of the video is part of the causal channel of protests for these specific cases. I partition these 30 cases into three groups: one, the first protest preceded the viral video; two, the viral video preceded (or concurrent) the first protest; three, no viral video. There was a significant reduction in lethal force following protests for all three groups, bolstering confidence in the baseline estimates. However, protests of high-profile police killings without viral videos did not reduce lethal force as much as protests with a viral video. It is unclear if this is because of the absence of a viral video, the small number of events, or a cohort effect.<sup>17</sup>

To further distinguish the impact of protesting from video recordings of police homicides, I report two falsification tests in Appendix D that leverage the timing and location of both videos and protests; if videos reduce lethal use-of-force in places without demonstrations, then I am likely overstating the impact of protests; likewise, if protests do not reduce lethal use-of-force without videos, then I am less confident in any direct protest effect. The results are mixed. As one may expect, places with protests and videos experienced significant

---

<sup>17</sup>High profile events without videos were more likely to occur in early cohorts (before the widespread distribution of police body-worn cameras). Only five of the thirty most prominent events did not have a video. Of these five events, three occurred in 2014, one in 2015, and one in 2016. On average, events without videos occur 1.6 years prior to the other 25 events. If the later cohorts of protests were more effective, then the result may be a cohort effect, not a video effect.

reductions in lethal force. Protests without videos also corresponded with a fall in lethal force, albeit less precise. In contrast, videos without protests corresponded with a rise in police homicides that is precise enough to rule out negative overall effect.

While it is unclear if protests or videos in-of-themselves drove the reduction, the evidence seems to favor protests. Of course, it is also possible that videos are heterogeneous, and the most compelling videos may have attracted protests. This argument cannot be dismissed given the relatively weaker results for protests without videos in both sets of tests. However, it is also possible that protests were heterogeneous. Suppose the most compelling demonstrations enhanced community vigilance or caused the local agency to adopt body-worn cameras or release body camera footage. In that case, future videos would select into places with effective protests. After all, protests preceded the viral videos in over half of the 30 most prominent sample police scandals, and Section IV.A associated protests with a significant increase in future body-worn cameras and video recordings of police homicides.

## V. Discussion

Stacked DID estimates suggest that census places with BLM protests experienced a 15% to 20% decrease in police homicides from 2014 to 2019, approximately 300 fewer police homicides. This fall in lethal use-of-force fell over time and became prominent when protests were large or frequent. While this reduction is robust to specification, estimator choice, choice of data, and population screens, it does not hold if lethal use-of-force is normalized by violent or total crime. The fall in lethal use-of-force may be partially explained by expansions in police body-worn cameras and community policing and depolicing.

Some caution is needed when interpreting these estimates. These DID estimates compare relatively large population cities to relatively smaller population places, implying that the control group is potentially inappropriate. This issue could not be resolved with inverse probability weighting. The variation in treatment timing is also limited. To be clear, it is entirely feasible that unobserved changes independent of protests between large and small cities drive these results. This issue is compounded by measurement error in lethal force data, which may correlate with population size and time; many larger cities have recently implemented online lethal force reporting; Fatal Encounters data is retroactive before 2013. However, the issue of measurement error is less severe; the results do not meaningfully change when using higher quality MPV data from 2013-2019. The results also hold when accounting for population-driven variance using per capita population-weighted least squares.

Simultaneity between viral videos and protests also warrants caution. BLM protests

often correspond with a high-profile police scandal, which garners public attention from both demonstrations and viral videos. The baseline estimates, therefore, may misattribute the impact of viral videos to protests. I alleviate this concern by showing that videos did not reduce lethal force when unaccompanied by demonstrations. I also show that lethal force fell after protests regardless of video status; however, the reduction for protests without videos was smaller and less precise. Furthermore, I show that protests preceded the viral video in over half of the 30 most prominent cases and the lethal force reduction was meaningful regardless of video status for these 30 cases.

There are several reasons to postulate that protests at least partially drove the rise in viral videos of police homicides. After demonstrations, local agencies were more likely to use body-worn cameras, making future recordings of fatal police encounters more likely. Protests also may have increased public scrutiny of the police, which may have emboldened onlookers to record police-civilian interactions. Finally, protests may have pressured police agencies to release footage of fatal police interactions.

The results indicate that civilian homicides increased by 10% following protests, exceeding the fall in lethal force due to the relative frequency. This estimate may tempt one into using a measure of lives saved/lost following protests to determine the social welfare implications of BLM. Such an analysis would not be convincing. The welfare implications for civilian and police homicides are distinct. The emerging literature on the spillover effects of police homicides makes this point clear. Police homicides do not diminish the tragedy of rising civilian homicides, but they too have a demonstrable impact on Black mental health (Bor et al., 2018) and future crime, including murders (Cheng and Long, 2018; Premkumar, 2020). They also profoundly threaten community trust and cooperation. Future research is needed to address this policy-relevant question with careful accounting outside the scope of this article.

## VI. Data Availability Statement

The data that support the findings of this study are openly available in “tcampbell3/Black-Lives-Matter-Lethal-Force: Black Lives Matter’s Effect on Police Lethal Use-of-Force Data and Replication Files” at <http://doi.org/10.5281/zenodo.4479837>, reference (Campbell, 2021). These data were derived from resources available in the public domain and are listed in Sections II.B.1-II.B.3.

## References

- Alpert, G. P. and Macdonald, J. M. (2001). Police use of force: An analysis of organizational characteristics. *Justice Quarterly*, (Issue 2):393.
- Ariel, B., Farrar, W., and Sutherland, A. (2015). The effect of police body-worn cameras on use of force and citizens' complaints against the police: A randomized controlled trial. *Journal of Quantitative Criminology*, 31(3):509 – 535.
- Ariel, B., Sutherland, A., Henstock, D., Young, J., Drover, P., Sykes, J., Megicks, S., and Henderson, R. (2016). Report: Increases in police use of force in the presence of body-worn cameras are driven by officer discretion: A protocol-based subgroup analysis of ten randomized experiments. *Journal of experimental criminology*, 12(3):453–463.
- Ariel, B., Sutherland, A., Henstock, D., Young, J., and Sosinski, G. (2018). The deterrence spectrum: Explaining why police body-worn cameras 'work' or 'backfire' in aggressive police-public encounters. *Policing: A Journal of Policy and Practice*, 12(1):6–26.
- Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2019). Synthetic difference in differences. Technical report, National Bureau of Economic Research.
- Ba, B. A., Knox, D., Mummolo, J., and Rivera, R. (2021). The role of officer race and gender in police-civilian interactions in chicago. *Science*, 371(6530):696–702.
- Ba, B. A. and Rivera, R. G. (2019). The effect of police oversight on crime and allegations of misconduct: Evidence from chicago.
- Baker, A., Larcker, D. F., and Wang, C. C. (2021). How much should we trust staggered difference-in-differences estimates? Available at SSRN 3794018.
- Banks, D., Couzens, L., Blanton, C., and Cribb, D. (2015). Arrest-related deaths program assessment: Technical report. *BJS Technical Reports*.
- Bor, J., Venkataramani, A., Williams, D., and Tsai., A. (2018). Police killings and their spillover effects on the mental health of black americans: a population-based, quasi-experimental study. *Lancet*, 10144(392):302–310.
- Braga, A. A., Sousa, W. H., Coldren, J. R., and Rodriguez, D. (2018). The effects of body-worn cameras on police activity and police-citizen encounters. *The Journal of Criminal Law and Criminology (1973-)*, 108(3):511–538.
- Buehler, J. W. (2017). Racial/ethnic disparities in the use of lethal force by us police, 2010–2014. *American journal of public health*, 107(2):295–297.
- Callaway, B. and Sant'Anna, P. H. (2020). Difference-in-differences with multiple time periods. *Journal of Econometrics*.
- Campbell, T. (2021). Black lives matter's effect on police lethal use-of-force data and replication files. Available at <http://doi.org/10.5281/zenodo.4479837>.
- Carleton, T. A. and Hsiang, S. M. (2016). Social and economic impacts of climate. *Science*, 353(6304).
- Cesario, J., Johnson, D. J., and Terrill, W. (2019). Is there evidence of racial disparity in police use of deadly force? analyses of officer-involved fatal shootings in 2015–2016. *Social Psychological and Personality Science*, 10(5):586–595.
- Cheng, C. and Long, W. (2018). The spillover effects of highly publicized police-related deaths on policing and crime: Evidence from large us cities. Technical report, Working Paper, July.

- Cunningham, J. P. and Gillezeau, R. (2018). Racial differences in police use of force: Evidence from the 1960s civil disturbances. *AEA Papers and Proceedings*, 108:217–21.
- 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.
- Donner, C. and Popovich, N. (2018). Hitting (or missing) the mark: An examination of police shooting accuracy in officer-involved shooting incidents. *Policing: An International Journal*.
- Einstein, K. L. and Kogan, V. (2016). Pushing the city limits: Policy responsiveness in municipal government. *Urban Affairs Review*, 52(1):3–32.
- Fryer, Roland G., J. (2019). An empirical analysis of racial differences in police use of force. *Journal of Political Economy*, 127(3):1210 – 1261.
- Fyfe, J. J. (2002). 'too many missing cases: Holes in our knowledge about police use of force.' *Justice Research & Policy*, 4(1/2):87 – 102.
- Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. Working Paper 25018, National Bureau of Economic Research.
- Hehman, E., Flake, J. K., and Calanchini, J. (2018). Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *SOCIAL PSYCHOLOGICAL AND PERSONALITY SCIENCE*, 9(4):393 – 401.
- Hoekstra, M. and Sloan, C. (2020). Does race matter for police use of force? evidence from 911 calls. Technical report, National Bureau of Economic Research.
- Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of economic literature*, 47(1):5–86.
- Jennings, W. G., Lynch, M. D., and Fridell, L. A. (2015). Evaluating the impact of police officer body-worn cameras (bwcs) on response-to-resistance and serious external complaints: Evidence from the orlando police department (opd) experience utilizing a randomized controlled experiment. *Journal of criminal justice*, 43(6):480–486.
- Joshua, C., Bernadette, P., Charles M., J., Bernd, W., Melody S., S., and Tracie, K. (2007). Across the thin blue line: Police officers and racial bias in the decision to shoot. *Journal of Personality and Social Psychology*, (6):1006.
- Kaplan, J. (2020). Jacob Kaplan's Concatenated Files: Uniform Crime Reporting Program Data: Offenses Known and Clearances by Arrest, 1960-2018.
- Kim, T. (2019). Facilitating police reform: Body cameras, use of force, and law enforcement outcomes. *Use of Force, and Law Enforcement Outcomes (October 23, 2019)*.
- Klinger, D., Rosenfeld, R., Isom, D., and Deckard, M. (2016). Race, crime, and the micro-ecology of deadly force. *Criminology & Public Policy*, 15(1):193–222.
- Kochel, T. R. (2019). Explaining racial differences in ferguson's impact on local residents' trust and perceived legitimacy: Policy implications for police. *Criminal Justice Policy Review*, 30(3):374–405.
- Legewie, J. and Fagan, J. (2016). Group threat, police officer diversity and the use of police force. *Columbia Public Law Research Paper*, 512(14).
- Long, W. (2019). How does oversight affect police? evidence from the police misconduct reform. *Journal of Economic Behavior & Organization*, 168:94–118.
- Madestam, A., Shoag, D., Veuger, S., and Yanagizawa-Drott, D. (2013). Do Political Protests Matter? Evidence from the Tea Party Movement\*. *The Quarterly Journal of Economics*, 128(4):1633–1685.

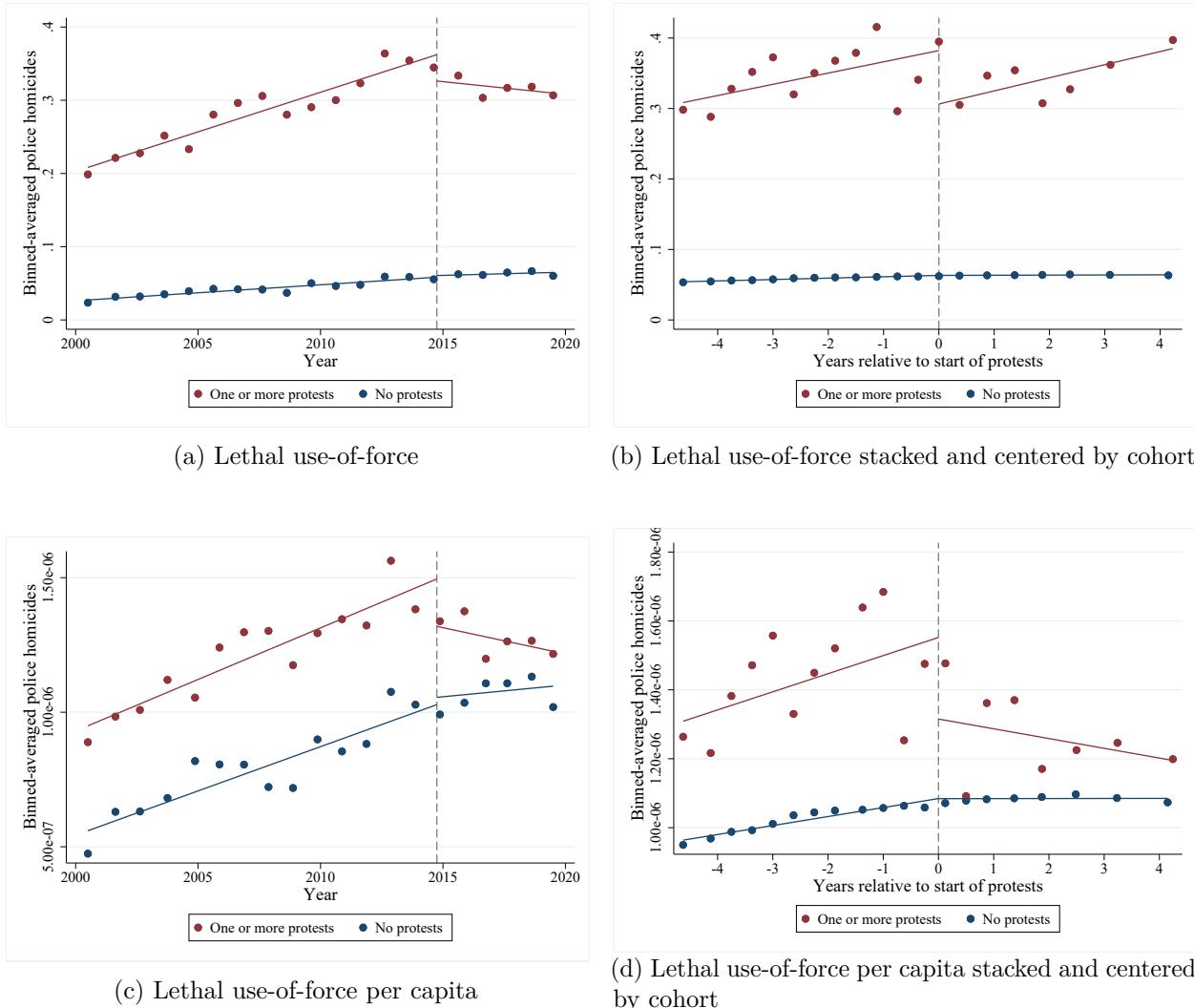
- Mazumder, S. (2019). Black lives matter for whites' racial prejudice: Assessing the role of social movements in shaping racial attitudes in the united states.
- McElvain, J. P. and Kposowa, A. J. (2008). Police officer characteristics and the likelihood of using deadly force. *Criminal Justice and Behavior*, (Issue 4):505.
- Menifield, C. E., Shin, G., and Strother, L. (2019). Do white law enforcement officers target minority suspects?. *Public Administration Review*, 79(1):56 – 68.
- 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. *Criminology and Public Policy*, (Issue 1):309.
- Paoline III, E. A. and Terrill, W. (2007). Police education, experience, and the use of force. *Criminal justice and behavior*, 34(2):179–196.
- Peterson, B. E., Yu, L., La Vigne, N., and Lawrence, D. S. (2018). The milwaukee police department's body-worn camera program. *Washington, DC: Urban Institute*.
- Premkumar, D. (2020). Intensified scrutiny and bureaucratic effort: Evidence from policing and crime after high-profile, officer-involved fatalities. *Officer-Involved Fatalities (October 19, 2020)*.
- Pyrooz, D., Decker, S., Wolfe, S., and Shjarback, J. (2016). Was there a ferguson effect on crime rates in large u.s. cities? *Journal of Criminal Justice*, 46:1–8.
- Ridgeway, G. (2020). The role of individual officer characteristics in police shootings. *The ANNALS of the American Academy of Political and Social Science*, 687(1):58–66.
- Ross, C. (2015). A multi-level bayesian analysis of racial bias in police shootings at the county-level in the united states, 2011-2014. *PLoS One.*, 11(10).
- Roth, J. (2019). Pre-test with caution: Event-study estimates after testing for parallel trends. *Unpublished manuscript, Department of Economics, Harvard University*.
- Sawyer, J. and Gampa, A. (2018). Implicit and explicit racial attitudes changed during black lives matter. *Personality and Social Psychology Bulletin*, 44(7):1039–1059. PMID: 29534647.
- Shjarback, J. A., Pyrooz, D. C., Wolfe, S. E., and Decker, S. H. (2017). De-policing and crime in the wake of ferguson: Racialized changes in the quantity and quality of policing among missouri police departments. *Journal of criminal justice*, 50:42–52.
- Skoy, E. (2020). Black lives matter protests, fatal police interactions, and crime. *Contemporary Economic Policy*.
- Solon, G., Haider, S. J., and Wooldridge, J. M. (2015). What are we weighting for? *Journal of Human Resources*, 50(2):301–316.
- Sun, L. and Abraham, S. (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*.
- Terrill, W. and Paoline, E. (2017). Police use of less lethal force: Does administrative policy matter? *Justice Quarterly*, 34(2):193–216.
- Tregle, B., Nix, J., and Alpert, G. P. (2019). Disparity does not mean bias: making sense of observed racial disparities in fatal officer-involved shootings with multiple benchmarks. *Journal of Crime and Justice*, 42(1):18–31.
- Trump, K.-S., Williamson, V., and Einstein, K. L. (2018). Vol 16(2): Replication Data for: Black Lives Matter: Evidence that Police- Caused Deaths Predict Protest Activity.
- United States. Bureau of Justice Statistics. (2018). Law enforcement agency identifiers crosswalk, united states, 2012. Available at <https://doi.org/10.3886/ICPSR35158.v2>.

- United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. (2015). Law Enforcement Management and Administrative Statistics (LEMAS), 2013.
- United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. (2019). Law Enforcement Management and Administrative Statistics Body-Worn Camera Supplement (LEMAS-BWCS), 2016.
- United States Department of Justice. Office of Justice Programs. Bureau of Justice Statistics. (2020). Law Enforcement Management and Administrative Statistics (LEMAS), 2016.
- U.S. Census Bureau (2013). Poverty Status in the PAST 12 Months 2013 American Community Survey 5-Year Estimates, Table S1701. Retrieved from <http://factfinder.census.gov>; (4 August 2019).
- U.S. Census Bureau (2019). SUB-EST00INT-TOT: Intercensal Estimates of the Resident Population for Incorporated Places and Minor Civil Divisions: April 1, 2000 to July 1, 2019. 2000-2019 Subcounty Intercensal Population Estimates. Retrieved from <https://www.census.gov/data/datasets/time-series/demo/popest/intercensal-2000-2010-cities-and-towns.html>; (28 July 2019) and <https://www2.census.gov/programs-surveys/popest/datasets/2010-2019/cities/totals/>; (6 June 2020).
- Van den Broek, T., Langley, D., and Hornig, T. (2017). The effect of online protests and firm responses on shareholder and consumer evaluation. *Journal of Business Ethics*, 146(2):279–294.
- Wallace, D., White, M. D., Gaub, J. E., and Todak, N. (2018). Body-worn cameras as a potential source of depolicing: Testing for camera-induced passivity. *Criminology*, 56(3):481–509.
- Wan, Y., Datta, S., Conklin, D., and Kong, M. (2015). Variable selection models based on multiple imputation with an application for predicting median effective dose and maximum effect. *Journal of Statistical Computation and Simulation*, 85(9):1902–1916.
- White, M. D. (2016). Transactional encounters, crisis-driven reform, and the potential for a national police deadly force database. *Criminology & Pub. Pol'y*, 15:223.
- White, M. D. and Malm, A. (2020). *Cops, cameras, and crisis: The potential and the perils of police body-worn cameras*. NYU Press.
- Yokum, D., Ravishankar, A., and Coppock, A. (2019). A randomized control trial evaluating the effects of police body-worn cameras. *Proceedings of the National Academy of Sciences*, 116(21):10329–10332.
- Zhang, T. H. (2016). Weather effects on social movements: Evidence from washington, d.c., and new york city, 1960–95. *Weather, Climate, and Society*, 8(3):299–311.
- Zhao, Y. and Long, Q. (2017). Variable selection in the presence of missing data: Imputation-based methods. *Wiley Interdisciplinary Reviews: Computational Statistics*, 9(5):e1402.

## List of Figures

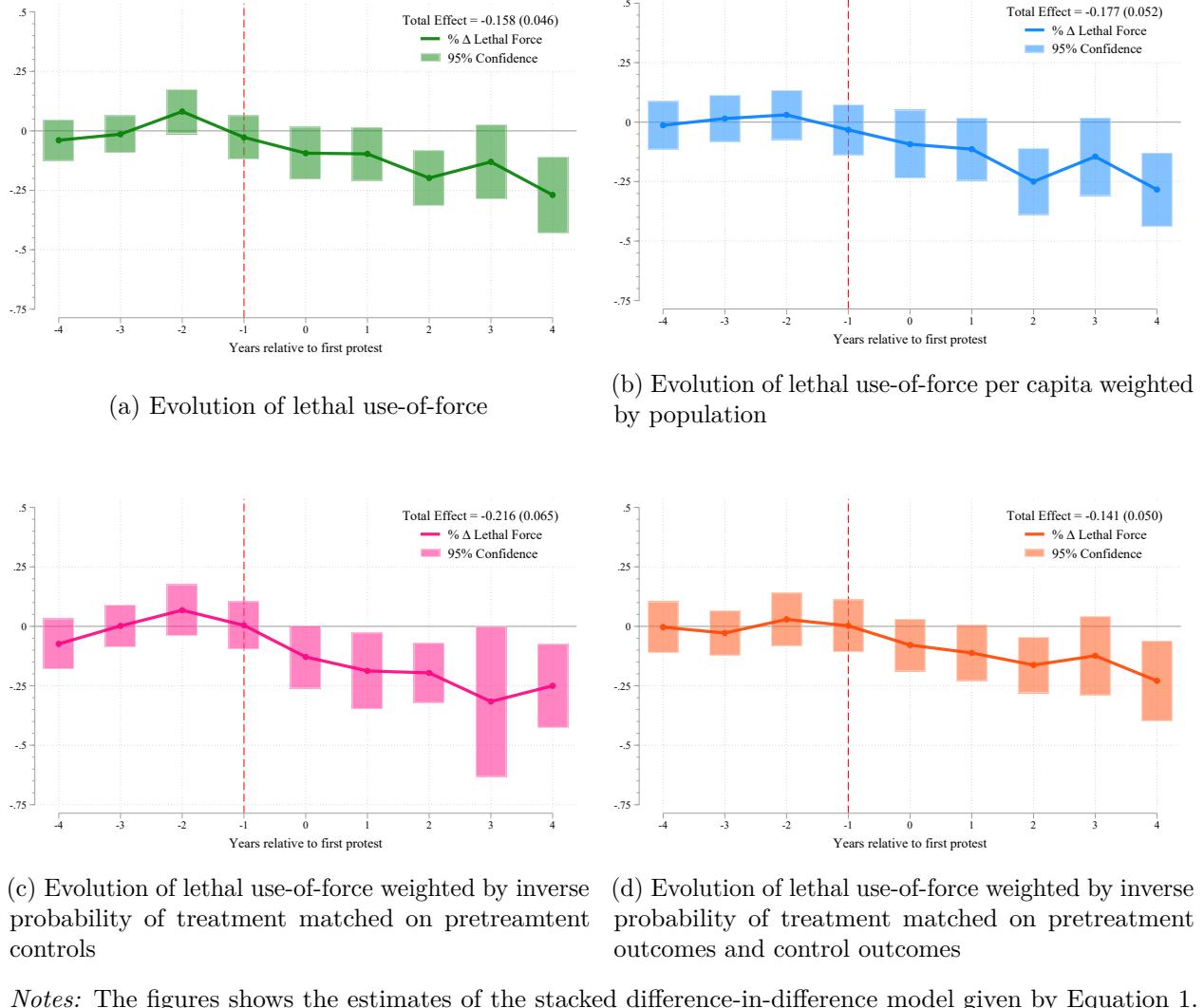
1	Binscatters of Lethal Use-of-force by Treatment Status . . . . .	29
2	Evolution of Impact of Black Lives Matter Protests on Police Homicides . . . . .	30
3	Evolution of the Share of Agencies with Body-Worn cameras (% $\Delta$ ) . . . . .	31
4	Differences in the Self-reported Reason Why Police Agencies Obtain Body-Worn Cameras by Protest Status . . . . .	32
5	Evolution of Impact of Black Lives Matter Protests on Crime . . . . .	33
A.1	Proportion of Total Fatal Encounters with the Police by Cause of Death from 2013-2019 . . . . .	38
A.2	Maps of Police Homicides and Black Lives Matter Protests . . . . .	39
A.3	Number of Cities with Black Lives Matter Protest Over Time . . . . .	40
A.4	Evolution of the Cumulative Number of Black Lives Matter Protests . . . . .	40
A.5	Robustness of Estimated Impact of Black Lives Matter Protests on Police Homicides to Specification . . . . .	41
A.6	Evolution of the Impact of Black Lives Matter Protests on Police Homicides with Videos . . . . .	42
B.1	Maps Black Lives Matter Protests Treatment Status by County . . . . .	51
B.2	Distribution of Synthetic Unit Weights over Counties . . . . .	52
B.3	Evolution of Impact of Black Lives Matter Protests on Police Homicides . . . . .	53
C.1	Heterogeneity in Black Lives Matter's Impact on Lethal Force from Video Timing Pooling Most Prominent Events . . . . .	57
C.2	Evolution of Video Timing Heterogeneity Estimates . . . . .	58
D.1	Evolution of Impact of Protests with Video of Police Homicides on Police Homicides . . . . .	61
D.2	Evolution of Impact of Video of Police Homicides with Protests on Other Police Homicides . . . . .	62
D.3	Evolution of Impact of Protests without Video of Police Homicides on Police Homicides . . . . .	63
D.4	Evolution of Impact of Video of Police Homicides without Protests on Other Police Homicides . . . . .	64

Figure 1: Binscatters of Lethal Use-of-force by Treatment Status



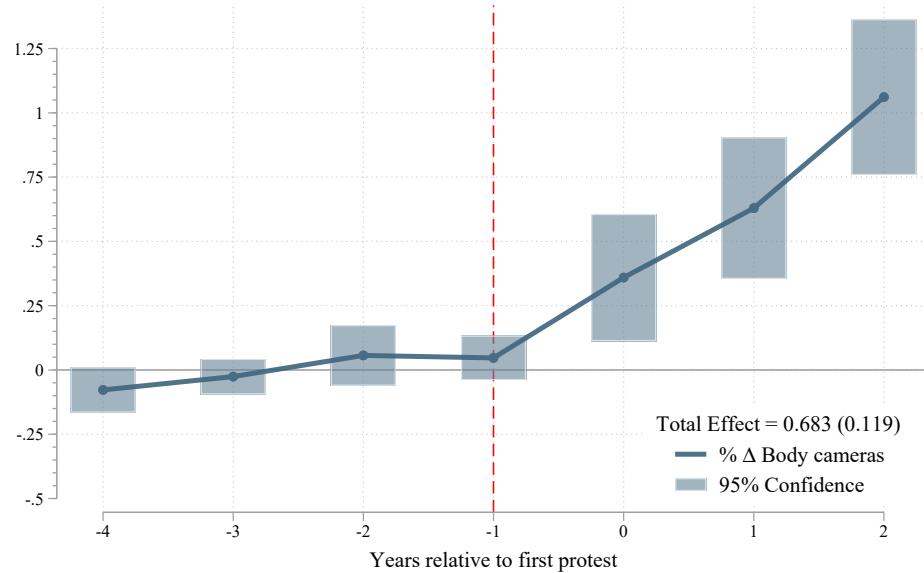
*Notes:* Figures 1a 1b report unweighted bin scatters of homicides by protest status. Figures 1c and 1d report population weighted bin scatter of homicides per capita by protest status. All figures also show the linear regression lines with a discontinuity at either 2014q3 or the quarter of the cohort's first protest aligned at 0.

Figure 2: Evolution of Impact of Black Lives Matter Protests on Police Homicides



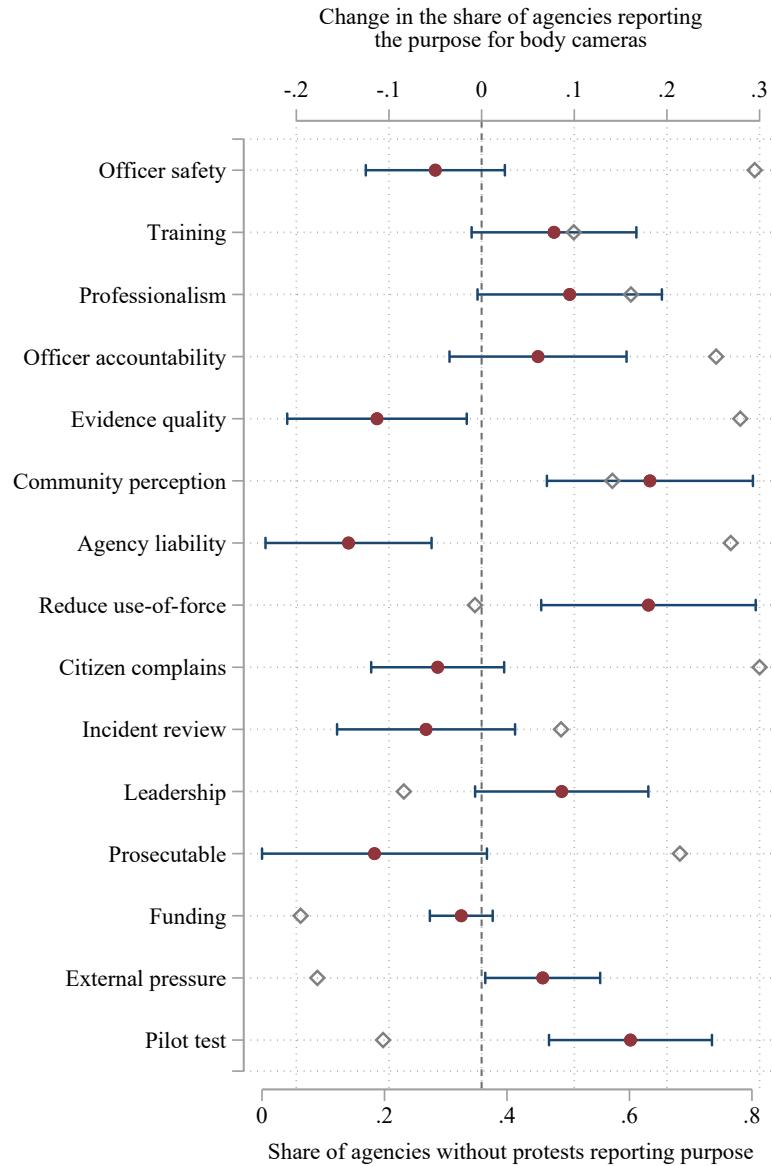
*Notes:* The figures show the estimates of the stacked difference-in-difference model given by Equation 1. Each sub-figure reports a different weighting scheme; all specifications include cohort-place and cohort-time fixed effects. The shaded area in each figure is the 95% confidence interval based on robust standard errors clustered by place. Figure 2a depicts the ordinary least squares estimates. Figure 2b displays the per capita weighted least squares (WLS) estimate that uses population weights to adjust for population-driven heteroscedasticity. Figure 2c displays the WLS estimate that uses lasso regularized inverse probability weights to balance pre-protest controls over protest exposure. Figure 2d displays the synthetic difference-in-differences estimates.

Figure 3: Evolution of the Share of Agencies with Body-Worn cameras (% $\Delta$ )



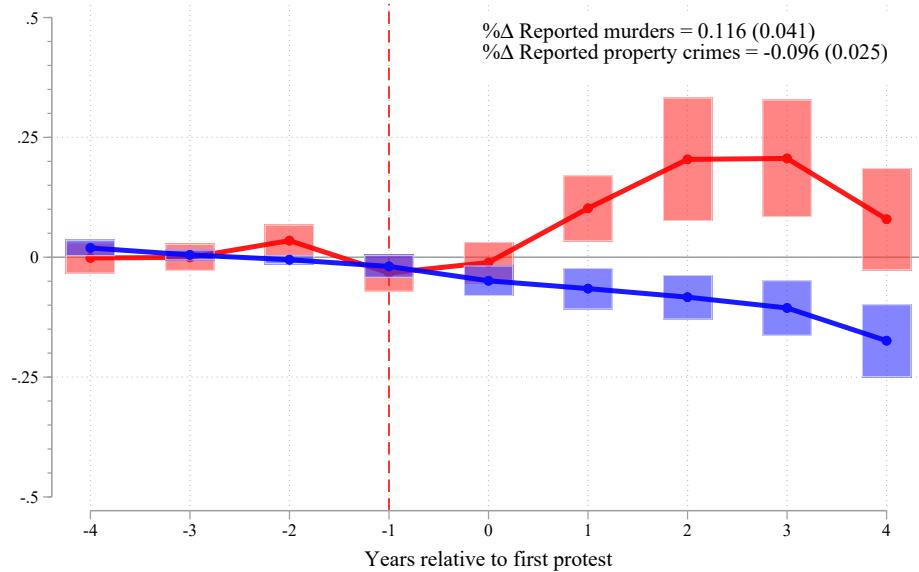
*Notes:* The figure shows estimates analogous to the stacked difference-in-difference model given by Equation 1 using an annual panel of police agencies. The regression controls for cohort-agency, cohort-time fixed effects, and a linear control for population interacted with cohort-population decile. The shaded area in the figure is the 95% confidence interval based on robust standard errors that are clustered by sampling unit and reported in parenthesis. Sampling weights are applied.

Figure 4: Differences in the Self-reported Reason Why Police Agencies Obtain Body-Worn Cameras by Protest Status

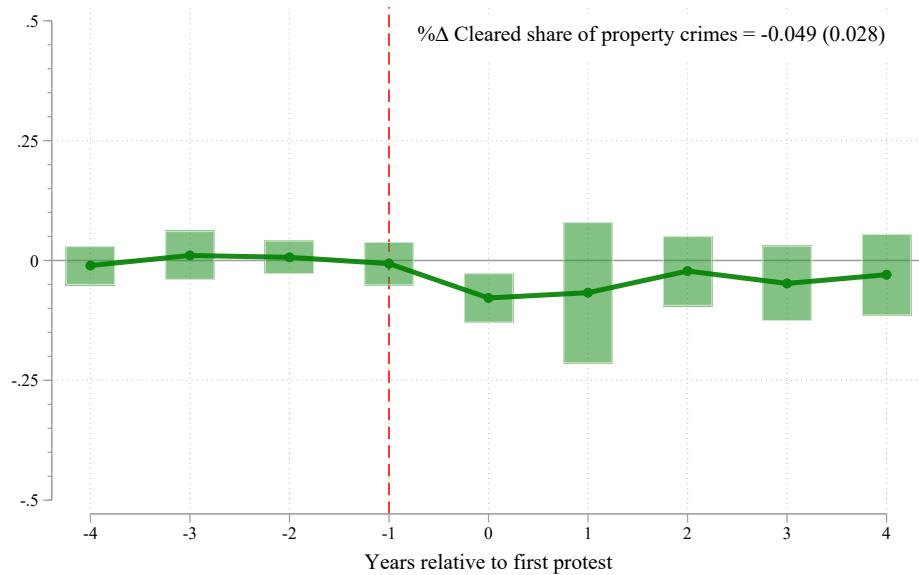


*Notes:* The figure depicts a sequence of cross-sectional regression estimates where the outcome is a dummy variable taking value one if the police agency reported the reason for obtaining body-worn cameras and zero if it is not a reason. The outcome is regressed on a treatment indicator, which equals one if the agency has at least one BLM protest from 2014-2016 and zero if the agency does not have any BLM protests from 2014-2016. The navy bars in the figure represent the 95% confidence interval based on robust standard errors that are clustered by sampling unit and reported in parenthesis. Sampling weights are applied.

Figure 5: Evolution of Impact of Black Lives Matter Protests on Crime



(a) Evolution of reported civilian homicides and property crimes



(b) Evolution of the share of total property crimes cleared by arrest

*Notes:* The figures shows estimates analogous to the stacked difference-in-difference model given by Equation 1 using an annual panel of police agencies. All regressions control for cohort-agency, cohort-time fixed effects, and a linear control for population interacted with cohort-population decile. The shaded area in each figure is the 95% confidence interval based on robust standard errors that are clustered by police agency. Red denotes the total reported murders, blue denotes total reported property crimes, and green denotes the ratio of property crimes cleared by arrest to total property crimes.

## List of Tables

1	Covariate Balance of Select Control Variables . . . . .	35
2	Impact of Black Lives Matter Protests on Police Homicides . . . . .	36
3	Impact of Black Lives Matter Protests on Police Agency Characteristics . . . . .	37
A.1	Effect of Protests on Police Homicides Using Mapping Police Violence Data .	43
A.2	Protest Effect on Police Homicides Heterogeneity by Size and Frequency . .	44
A.3	Robustness of Estimates to Normalization . . . . .	45
A.4	Robustness of Estimates to Different Population Screens . . . . .	46
A.5	Robustness of Estimates to Regression Specification . . . . .	46
A.6	Impact of Black Lives Matter Protests on Crime . . . . .	47
C.1	List of High Profile Police Killings from 2014 to 2019 . . . . .	56

Table 1: Covariate Balance of Select Control Variables

	Unweighted		Population		IPW Controls		IPW Unit		IPW Unit-Time	
	(1) Treated	(2) Control	(3) Treated	(4) Control	(5) Treated	(6) Control	(7) Treated	(8) Control	(9) Treated	(10) Control
Poverty	18.72 (4.98)	21.11 (4.50)	18.81 (3.66)	21.39 (4.26)	18.72 (4.98)	19.31 (4.47)	18.72 (4.98)	21.45 (4.24)	18.40 (4.97)	21.51 (4.18)
Labor force participation rate	64.42 (5.50)	65.48 (6.38)	65.68 (4.14)	65.90 (5.86)	64.42 (5.50)	64.02 (6.16)	64.42 (5.50)	65.60 (5.82)	64.50 (5.42)	65.62 (5.71)
Unemployment rate	10.52 (3.49)	10.02 (3.69)	10.99 (3.15)	10.14 (3.53)	10.52 (3.49)	10.76 (4.11)	10.52 (3.49)	10.28 (3.53)	10.53 (3.55)	10.34 (3.50)
Full time employment rate	2.72 (0.58)	3.15 (0.74)	2.79 (0.37)	3.14 (0.66)	2.72 (0.58)	2.77 (0.54)	2.72 (0.58)	3.15 (0.68)	2.68 (0.56)	3.15 (0.67)
Black population	19.04 (17.15)	11.30 (16.22)	21.68 (16.38)	10.87 (15.30)	19.04 (17.15)	18.05 (21.82)	19.04 (17.15)	11.08 (15.99)	20.01 (17.29)	11.07 (16.01)
Black poverty rate	28.73 (14.43)	22.85 (17.08)	30.17 (9.00)	22.58 (14.55)	28.73 (14.43)	25.44 (16.75)	28.73 (14.43)	22.69 (14.66)	29.15 (14.13)	22.66 (14.19)
< High school	13.57 (6.71)	13.76 (9.12)	16.54 (5.88)	14.49 (9.31)	13.57 (6.71)	12.88 (7.43)	13.57 (6.71)	15.11 (9.67)	13.64 (6.68)	15.34 (9.74)
High school	24.08 (7.58)	26.23 (7.99)	23.71 (5.58)	25.48 (7.45)	24.08 (7.58)	25.15 (8.76)	24.08 (7.58)	25.62 (7.31)	23.71 (7.52)	25.52 (7.18)
Some college	21.62 (4.43)	23.33 (4.57)	20.58 (4.09)	23.40 (4.44)	21.62 (4.43)	21.95 (4.44)	21.62 (4.43)	23.45 (4.40)	21.28 (4.55)	23.47 (4.37)
College	34.39 (15.29)	29.82 (15.07)	32.88 (10.65)	29.55 (14.32)	34.39 (15.29)	33.53 (16.98)	34.39 (15.29)	28.75 (14.12)	35.13 (15.33)	28.55 (13.92)
Population (100,000s)	2.43 (5.91)	0.57 (0.44)	16.76 (25.45)	0.90 (0.71)	2.43 (5.91)	1.07 (1.02)	2.43 (5.91)	1.03 (0.82)	2.77 (6.53)	1.12 (0.85)
Officer safety	0.14 (0.24)	0.31 (0.42)	0.13 (0.22)	0.28 (0.39)	0.14 (0.24)	0.16 (0.28)	0.14 (0.24)	0.28 (0.38)	0.13 (0.23)	0.27 (0.38)
Violent crime rate	0.01 (0.00)	0.00 (0.00)								
Property crime rate	0.04 (0.02)	0.03 (0.02)	0.04 (0.02)	0.03 (0.02)	0.04 (0.02)	0.04 (0.02)	0.04 (0.02)	0.03 (0.02)	0.04 (0.02)	0.03 (0.01)
Officer wage	24.93 (8.82)	27.52 (9.48)	23.00 (6.51)	27.68 (9.42)	24.93 (8.82)	25.40 (8.29)	24.93 (8.82)	27.62 (8.86)	24.66 (8.75)	27.64 (8.75)
Share of black officers	0.12 (0.12)	0.06 (0.11)	0.15 (0.13)	0.06 (0.09)	0.12 (0.12)	0.10 (0.13)	0.12 (0.12)	0.06 (0.08)	0.12 (0.12)	0.05 (0.08)
Population density (10,000s per mile)	0.38 (0.31)	0.36 (0.37)	0.74 (0.79)	0.39 (0.37)	0.38 (0.31)	0.37 (0.42)	0.38 (0.31)	0.38 (0.34)	0.40 (0.33)	0.38 (0.34)
2008 pres. democratic vote share	0.64 (0.15)	0.56 (0.15)	0.69 (0.14)	0.56 (0.15)	0.64 (0.15)	0.61 (0.15)	0.64 (0.15)	0.56 (0.14)	0.66 (0.15)	0.56 (0.14)
Observations	314	1257	314	1257	314	1257	314	1257	314	1257
F <sub>18, 26902</sub>	44.00	44.00	34.38	34.38	2.260	2.260	38.29	38.29	41.26	41.26

*Notes:* This table displays the 2013 average values for places that eventually have a Black Lives Matters protest (treated) and for places that do not have a Black Lives Matter protests during the sample (Control). The column titles refer to different weights described in Section II. Population refers to weighting by the annual population, IPW Controls refers to weighting by the inverse probability of eventually having a protest using pre-BLM control variables, IPW Unit refers to weighting by the inverse probability of eventually having a protest using annual incidents of lethal force as covariates, and IPW Unit-Time refers to weighting by the product of IPW unit and IPW time weights. The row F-stat reports the joint F-test for difference in means based on robust standard errors.

Table 2: Impact of Black Lives Matter Protests on Police Homicides

	(1)	(2)	(3)	(4)	(5)	(6)
%ΔLethal Force	-0.158 (0.046)	-0.177 (0.052)	-0.216 (0.065)	-0.143 (0.052)	-0.154 (0.044)	-0.141 (0.050)
ΔTotal Lethal Force	288 (83.9)	323 (94.8)	393 (119)	261 (94.8)	313 (89.8)	288 (102)
Average outcome pre-protest ( $\bar{y}_{N-1}$ )	0.347	0.000	0.347	0.347	0.389	0.389
Average normalization pre-protest ( $\bar{N}_{-1}$ )	1	245,080	1	1	1	1
Total place-quarters after protest ( $e$ )	5252	5252	5252	5252	5252	5252
Total lethal force post-protest	1,847	1,847	1,847	1,847	1,847	1,847
Places with protests	314	314	314	314	314	314
Places without protests	1,257	1,257	1,257	1,257	1,257	1,257
Total number of protests	1,724	1,724	1,724	1,724	1,724	1,724
Total number of protesters	347,133	347,133	347,133	347,133	347,133	347,133
Number of cohorts	22	22	22	22	22	22
Sample size	2,219,276	2,219,276	2,219,276	2,219,276	2,219,276	2,219,276
Normalization	None	Population	None	None	None	None
Population weights		✓				
Pre-treatment control inverse probability weights			✓			
Event-place inverse probability weights				✓		✓
Event-quarter inverse probability weights					✓	✓

*Notes:* This table reports the benchmark estimates detailed in Section II. All regressions control for population decile, a linear population-population decile interaction, cohort-census place, and cohort-time (quarterly) fixed effects. Standard errors are clustered by census place and reported in parenthesis.

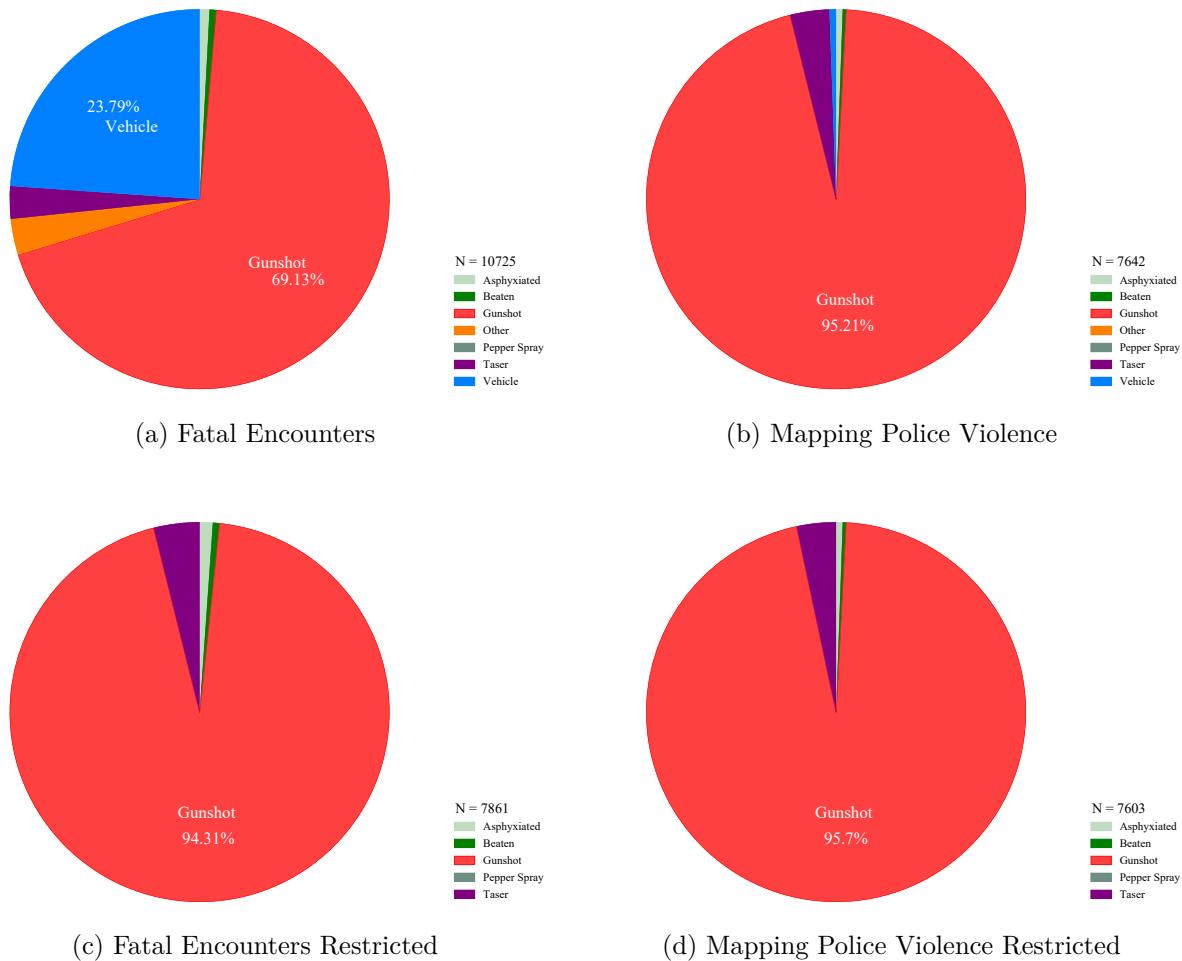
Table 3: Impact of Black Lives Matter Protests on Police Agency Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Outcome	Body cameras	Patrol officers	Sara officers	Black officers	White officers	Exp. officers	College required	Force doc.	Budget (millions)
Impact of protest (% $\Delta$ )	0.683 (0.119)	0.206 (0.136)	0.575 (0.326)	-0.067 (0.033)	-0.011 (0.011)	-0.022 (0.014)	-0.348 (0.268)	0.056 (0.043)	0.030 (0.033)
Average outcome pre-protest ( $\bar{y}_{N-1}$ )	0.238	203	96.3	156	547	867	0.132	0.968	134
Agencies with protests	220	197	197	197	197	197	197	197	197
Agencies without protests	3,634	798	798	798	798	798	798	798	798
Total number of protests	1,235	1,103	1,103	1,103	1,103	1,103	1,103	1,103	1,103
Total number of protesters	309,423	282,030	282,030	282,030	282,030	282,030	282,030	282,030	282,030
Sample size	2,070,116	1,818	1,839	1,876	1,876	1,850	1,885	1,909	1,860
Years	2000-2016	2013, 2016	2013, 2016	2013, 2016	2013, 2016	2013, 2016	2013, 2016	2013, 2016	2013, 2016
Time unit	Quarter	Annual	Annual	Annual	Annual	Annual	Annual	Annual	Annual

*Notes:* This table reports the impact of Black Lives Matter protests on various police agency characteristics. Data comes from the Law Enforcement Administrative Statistics (LEMAS) 2013, 2016, and 2016 Body-Worn Camera supplement. All regressions control for cohort-agency, cohort-time fixed effects, and a linear control for population interacted with cohort-population decile. Standard errors are clustered by sampling unit and reported in parenthesis. Sampling weights are applied. Body-worn cameras indicate the agency obtained body-worn cameras, Patrol officers are the number of officers with designated geographic patrol areas, SARA officers are the number of officers encouraged to engage in SARA-type problem-solving, Black officers is the number of Black officers, White officers is the number of White officers, Exp. officers are the number of officers less new recruits during the survey year, College required indicates a two-year college degree requirement for hires, and Force doc. indicates the agency requires documentation for the following types of force: chemical, gun discharge, gun display, or neck restraint. Budget is the agency's annual operating budget in millions of dollars.

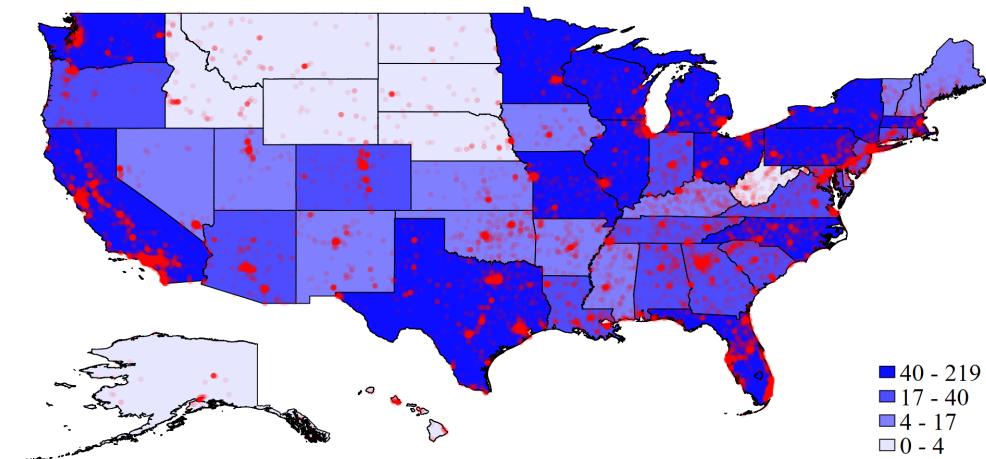
## A. Appendix – Additional Tables and Figures

Figure A.1: Proportion of Total Fatal Encounters with the Police by Cause of Death from 2013-2019

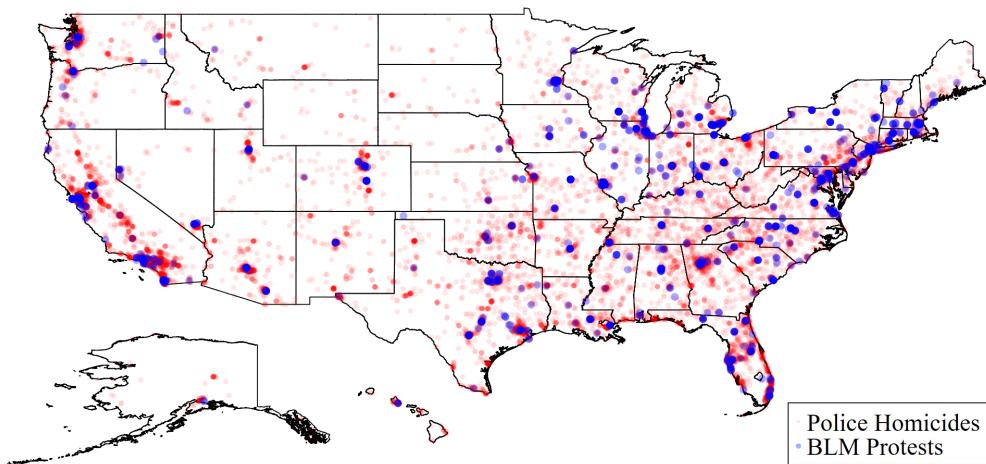


*Notes:* The figure shows the proportion of fatal encounters with police by causes of death from 2013-2019. The figure compares the Fatal Encounters and Mapping Police Violence datasets.

Figure A.2: Maps of Police Homicides and Black Lives Matter Protests



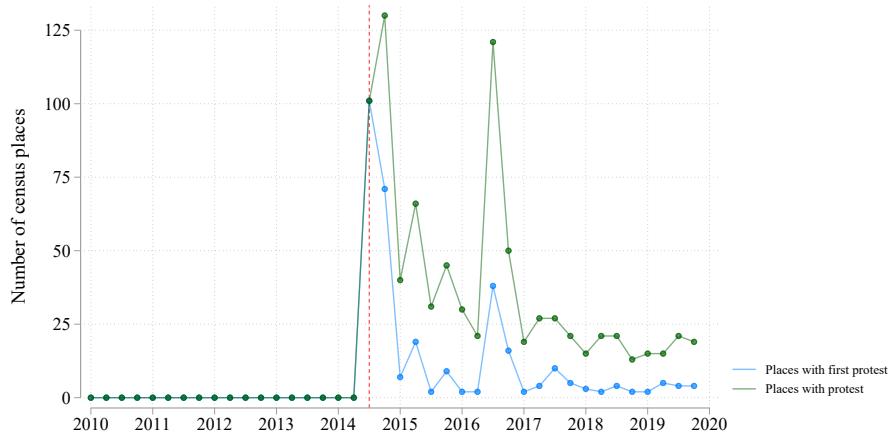
(a) Police Homicides over Protest Totals by State



(b) Police Homicides and Daily Protests

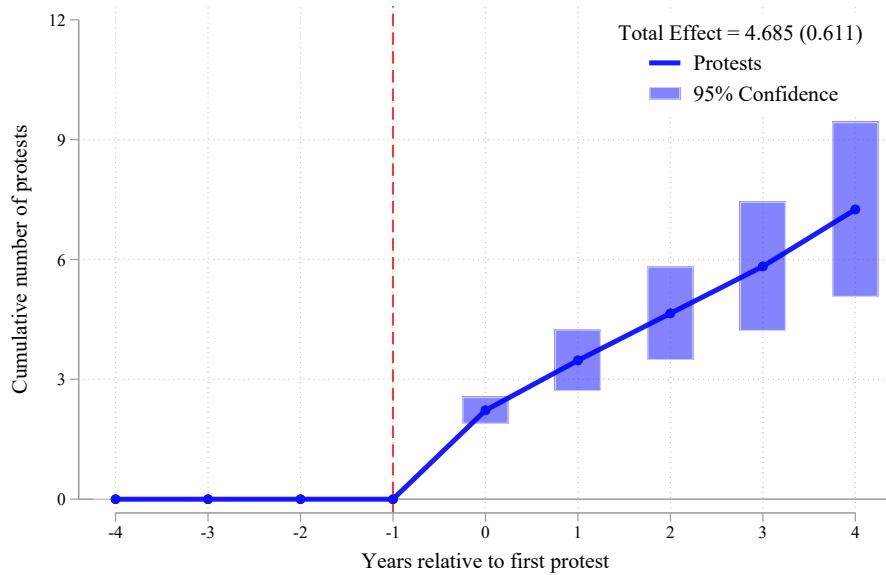
*Notes:* The figures show the location of police homicides and Black Lives Matter (BLM) protests from 2000 to 2019. Blue denotes BLM protests. Red indicates a police homicide.

Figure A.3: Number of Cities with Black Lives Matter Protest Over Time



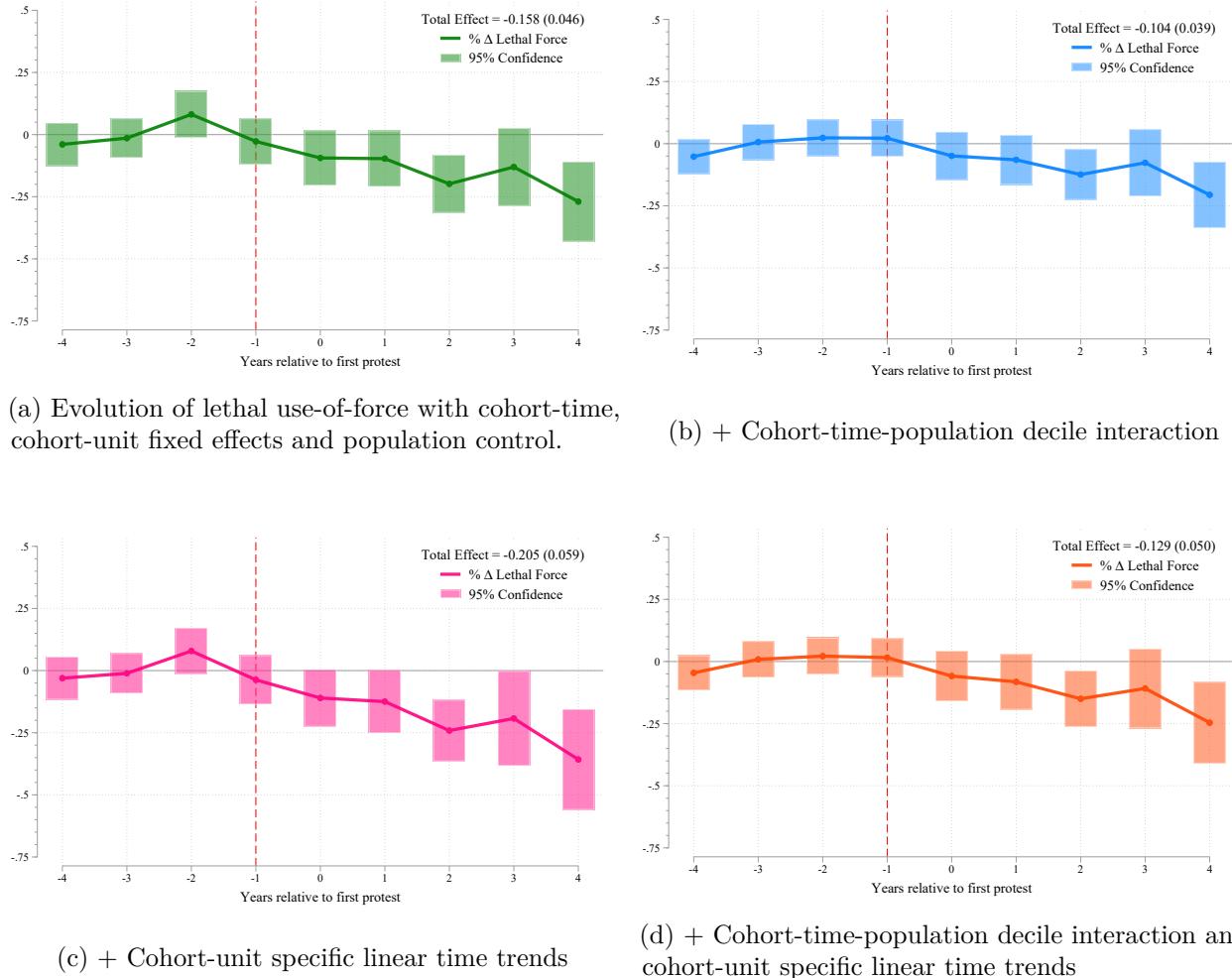
Notes: The figure shows the number of census places that experience their first Black Lives Matter protests during each quarter from 2010 through 2019 in blue. The total number of cities with Black Lives Matter protests during the quarter is displayed in green. Census places with a population under 20,000 at any point during the sample are omitted.

Figure A.4: Evolution of the Cumulative Number of Black Lives Matter Protests



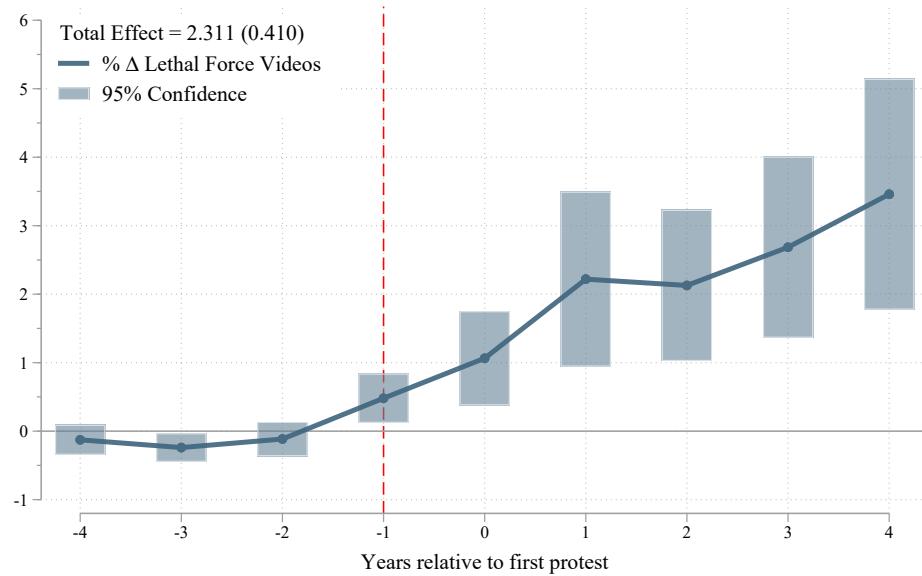
Notes: The figure shows the annual, average difference in the cumulative number of Black Lives Matter (BLM) protests between the treated and control census places. The blue line gives the  $\hat{\beta}_k$  from a regression of the total number of BLM protests on the right-hand side of Equation 1. The regression includes cohort-time and cohort-place fixed effects. The blue shaded area is the 95% confidence interval based on robust standard errors that are clustered by place.

Figure A.5: Robustness of Estimated Impact of Black Lives Matter Protests on Police Homicides to Specification



*Notes:* The figure assess the robustness the stacked difference-in-difference model given by Equation 1 to alternative specifications. All specifications include cohort-place, cohort-time, cohort-population decile fixed effects and are estimated with ordinary least squares. The shaded area in each figure is the 95% confidence interval based on robust standard errors that are clustered by place. Figure A.5a depicts the benchmark estimates. Figure A.5b shows estimates that also include an a cohort-time-population decile interaction. Figure A.5c displays estimates that also include cohort-unit specific linear time trends. Figure A.5d displays estimates that also include both a cohort-time-population decile interaction and unit specific linear time trends.

Figure A.6: Evolution of the Impact of Black Lives Matter Protests on Police Homicides with Videos



*Notes:* The figure shows the annual, average difference in the number of police homicides with recorded by a body camera between the treated and control census places. The blue line gives the  $\hat{\beta}_k$  from a regression of Equation 1. The regression includes cohort-time and cohort-place fixed effects along with a linear population-population decile interaction. The shaded area is the 95% confidence interval based on robust standard errors that are clustered by place. The outcome includes 4 bystander videos, 3 dashcam videos, 7 surveillance videos, and 601 body camera videos. An under-count of videos is almost certain prior to 2015; the year when the Washington Post began document police shootings with body-worn cameras.

Table A.1: Effect of Protests on Police Homicides Using Mapping Police Violence Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
%ΔLethal Force	-0.134 (0.048)	-0.153 (0.061)	-0.151 (0.093)	-0.106 (0.086)	-0.131 (0.088)	-0.146 (0.117)	-0.175 (0.094)	-0.235 (0.120)
ΔTotal Lethal Force	241 (86.5)	275 (110)	88 (54.1)	75 (61.3)	92 (61.5)	104 (83.4)	90 (48.3)	121 (61.7)
Average outcome pre-protest ( $\bar{Y}_{N-1}$ )	0.343	0.000	0.111	0.000	0.133	0.000	0.098	0.000
Average normalization pre-protest ( $\bar{N}_{-1}$ )	1	245080	1	141521	1	53746	1	245080
Total place-quarters after protest ( $e$ )	5252	5252	5252	5252	5252	5252	5252	5252
Total lethal force post-protest	1,769	1,769	515	515	728	728	488	488
Places with protests	314	314	314	314	314	314	314	314
Places without protests	1,257	1,257	1,257	1,257	1,257	1,257	1,257	1,257
Total number of protests	1,724	1,724	1,724	1,724	1,724	1,724	1,724	1,724
Total number of protesters	347,133	347,133	347,133	347,133	347,133	347,133	347,133	347,133
Sample size	487,310	487,310	378,326	378,326	378,326	378,326	487,310	487,310
Police homicide subset	Total	Total	White	White	Black	Black	Unarmed	Unarmed
Benchmark	None	Popula-tion	None	White	None	Black	None	Popula-tion
Weight	None	Popula-tion	None	White	None	Black	None	Popula-tion

*Notes:* This table reports the stacked difference-in-differences estimates using different data from Mapping Police Violence and decomposes incidents of lethal force by race or alleged arming of the victim. All regressions control for cohort-population decile, a linear population-cohort-population decile interaction, cohort-census place, and cohort-time (quarterly) fixed effects. Standard errors are clustered by census place and reported in parenthesis.

Table A.2: Protest Effect on Police Homicides Heterogeneity by Size and Frequency

	(1)	(2)	(3)	(4)	(5)
Maximum protest size					
Quartile 1 ( $\leq 30$ )	-0.064 (0.116)	-0.081 (0.119)	-0.084 (0.118)	-0.182 (0.133)	-0.151 (0.135)
Quartile 2 ( $\leq 100, > 30$ )	-0.039 (0.084)	-0.069 (0.083)	-0.086 (0.083)	-0.140 (0.098)	-0.116 (0.106)
Quartile 3 ( $\leq 300, > 100$ )	-0.009 (0.081)	-0.060 (0.087)	-0.032 (0.084)	-0.128 (0.137)	-0.094 (0.138)
Quartile 4 ( $> 300$ )	-0.163 (0.070)	-0.219 (0.065)	-0.211 (0.066)	-0.252 (0.093)	-0.195 (0.080)
Total number of protests					
Quartile 1 ( $\leq 1$ )	0.006 (0.101)	-0.004 (0.102)	-0.008 (0.103)	-0.050 (0.122)	-0.030 (0.124)
Quartile 2 ( $\leq 2, > 1$ )	-0.187 (0.138)	-0.224 (0.141)	-0.226 (0.140)	-0.355 (0.163)	-0.324 (0.169)
Quartile 3 ( $\leq 4, > 2$ )	0.106 (0.142)	0.077 (0.140)	0.080 (0.140)	-0.033 (0.162)	0.021 (0.178)
Quartile 4 ( $> 4$ )	-0.143 (0.055)	-0.190 (0.052)	-0.184 (0.052)	-0.228 (0.077)	-0.186 (0.068)
Cohort-place fixed effects	✓	✓	✓	✓	✓
Cohort-time fixed effects	✓	✓	✓	✓	✓
Population controls		✓	✓	✓	✓
Consent decree controls			✓	✓	✓
Cohort-place linear time trend				✓	✓
Cohort-time-population fixed effects					✓

*Notes:* This table assess heterogeneity in the impact of protests on police homicides by the size and frequency of protests using the same specification given by Equation 3, except I am running a separate regression by maximum protest size quartile or the total number of protest quartile. To be clear, each regression omits events that are outside of the indicated protest size or frequency quartile. However, control cities are always included; meaning, the reference category is always the control group, cities that do not have protests during the sample. All regressions control for population decile, a linear population-population decile interaction, cohort-census place, and cohort-time (quarterly) fixed effects. Standard errors are clustered by census place and reported in parenthesis.

Table A.3: Robustness of Estimates to Normalization

	(1)	(2)	(3)	(4)	(5)
%ΔLethal Force	-0.158 (0.046)	-0.095 (0.071)	-0.347 (0.259)	12.093 (9.076)	16.784 (12.833)
ΔTotal Lethal Force	288 (83.9)	176 (131)	932 (697)	-42,938 (32,226)	-63,817 (48,793)
Average outcome pre-protest ( $\bar{Y}_{N-1}$ )	0.347	0.000	0.001	288	79.1
Average normalization pre-protest ( $\bar{N}_{-1}$ )	1	245,080	696	0.002	0.009
Total place-quarters after protest ( $e$ )	5252	5252	5252	5252	5252
Total lethal force post-protest	1,847	1,847	1,847	1,847	1,847
Places with protests	314	314	314	314	314
Places without protests	1,257	1,257	1,257	1,257	1,257
Total number of protests	1,724	1,724	1,724	1,724	1,724
Total number of protesters	347,133	347,133	347,133	347,133	347,133
Benchmark	None	Popula-tion	Officers	Violent Arrests	Total Arrests
Sample size	2,219,276	2,219,276	1,342,692	1,929,804	1,946,905

*Notes:* This table reports the robustness of the estimates to using various benchmark variables (dividing policing homicides by a variable prior to the regression). The baseline specification is Equation 3. All regressions control for population decile, a linear population-population decile interaction, cohort-census place, and cohort-time (quarterly) fixed effects. Standard errors are clustered by census place and reported in parenthesis.

Table A.4: Robustness of Estimates to Different Population Screens

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
% $\Delta$ Lethal Force	-0.158 (0.046)	-0.161 (0.046)	-0.156 (0.047)	-0.150 (0.047)	-0.146 (0.049)	-0.187 (0.059)	-0.165 (0.090)
$\Delta$ Total Lethal Force	288 (83.9)	288 (82.4)	271 (81.8)	266 (83.3)	260 (87.2)	312 (98.4)	222 (121)
Average outcome pre-protest ( $\bar{y}_{N-1}$ )	0.347	0.422	0.519	0.596	0.696	0.962	1.153
Average normalization pre-protest ( $\bar{N}_{-1}$ )	1	1	1	1	1	1	1
Total place-quarters after protest ( $e$ )	5252	4243	3352	2976	2556	1733	1170
Total lethal force post-protest	1,847	1,798	1,736	1,692	1,657	1,513	1,306
Places with protests	314	248	195	167	140	92	62
Places without protests	1,257	547	285	166	97	25	5
Total number of protests	1,724	1,589	1,497	1,444	1,377	1,219	1,091
Total number of protesters	347,133	330,450	321,530	317,818	310,781	291,149	274,939
Population screen	20,000	40,000	60,000	80,000	100,000	175,000	250,000
Number of cohorts	22	22	21	18	14	9	7
Sample size	2,219,276	973,598	489,934	249,508	118,770	25,111	7,611

*Notes:* This table reports the robustness of the estimates to using various population screens (omitting observations with a population below the screen at any time during the sample). The baseline specification is Equation 3. All regressions control for population decile, a linear population-population decile interaction, cohort-census place, and cohort-time (quarterly) fixed effects. Standard errors are clustered by census place and reported in parenthesis.

Table A.5: Robustness of Estimates to Regression Specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
% $\Delta$ Lethal Force	-0.107 (0.049)	-0.085 (0.041)	-0.229 (0.068)	-0.158 (0.046)	-0.159 (0.046)	-0.111 (0.051)	-0.103 (0.049)	-0.257 (0.088)	-0.112 (0.066)	-0.095 (0.062)
$\Delta$ Total Lethal Force	195 (89.3)	155 (74.7)	417 (124)	288 (83.9)	291 (83.9)	203 (93.0)	187 (89.3)	469 (160)	204 (120)	194 (127)
Average outcome pre-protest ( $\bar{y}_{N-1}$ )	0.347	0.347	0.347	0.347	0.347	0.347	0.347	0.347	0.347	0.389
Average normalization pre-protest ( $\bar{N}_{-1}$ )	1	1	1	1	1	1	1	1	1	1
Total place-quarters after protest ( $e$ )	5252	5252	5252	5252	5252	5252	5252	5252	5252	5252
Total lethal force post-protest	1,847	1,847	1,847	1,847	1,847	1,847	1,847	1,847	1,847	1,847
Places with protests	314	314	314	314	314	314	314	314	314	314
Places without protests	1,257	1,257	1,257	1,257	1,257	1,257	1,257	1,257	1,257	1,257
Total number of protests	1,724	1,724	1,724	1,724	1,724	1,724	1,724	1,724	1,724	1,724
Total number of protesters	347,133	347,133	347,133	347,133	347,133	347,133	347,133	347,133	347,133	347,133
Number of cohorts	22	22	22	22	22	22	22	22	22	22
Sample size	2,219,276	2,219,276	2,219,276	2,219,276	2,219,276	779,312	750,716	750,716	750,716	750,716
Cohort-place fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort-time fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort-time-population quintile fixed effects										
Cohort-place linear time trend		✓								
Population controls				✓	✓	✓	✓	✓	✓	✓
Consent decree controls					✓	✓	✓	✓	✓	✓
Demographic and labor market controls						✓	✓	✓	✓	✓
Crime controls							✓	✓	✓	✓
Pre-treatment control inverse probability weights								✓		
Event-place inverse probability weights									✓	
Event-place and event-quarter inverse probability weights										✓

*Notes:* This table reports the robustness of the estimates to using various benchmark specifications of time variant control variables, inverse probability weights, and fixed effects. Standard errors are clustered by census place and reported in parenthesis.

Table A.6: Impact of Black Lives Matter Protests on Crime

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome	Total murders	Total violent crimes	Cleared violent crimes	Total property crimes	Cleared property crimes	Share of property crimes cleared	Officer assaults
Impact of protest (% $\Delta$ )	0.116 (0.041)	0.030 (0.027)	0.065 (0.099)	-0.096 (0.025)	-0.126 (0.052)	-0.049 (0.028)	-0.020 (0.076)
Average outcome pre-protest ( $\bar{y}_{N-1}$ )	7.99	626	249	3,393	507	0.162	17.1
Agencies with protests	963	963	963	963	963	963	963
Agencies without protests	21,056	21,056	21,056	21,056	21,056	21,056	21,056
Total number of protests	6,453	6,453	6,453	6,453	6,453	6,453	6,453
Total number of protesters	1,244,639	1,244,639	1,244,639	1,244,639	1,244,639	1,244,639	1,244,639
Sample size	1,826,160	1,826,147	1,826,153	1,826,035	1,826,035	1,615,031	1,826,160
Years	2000-2019	2000-2019	2000-2019	2000-2019	2000-2019	2000-2019	2000-2019
Time unit	Annual	Annual	Annual	Annual	Annual	Annual	Annual

*Notes:* This table reports the impact of Black Lives Matter protests on crime counts. All regressions control for cohort-agency, cohort-time fixed effects, and a linear control for population interacted with cohort-population decile. Standard errors are clustered by agency and reported in parenthesis.

## B. Appendix – County Level Analysis to Account for Local Spillovers and Invalid Control Units.

The baseline estimate shows that census places experience a 15%-20% decrease in police homicides following protests relative to generally less populous places without protests. This result holds when inverse probability weights are used to balance lethal force trends between treated and control places. One may be concerned that these weights may be picking up noise in the lethal use-of-force measure and, therefore, may emphasize the wrong control units. Furthermore, one may be concerned that the baseline estimates fail to account for spillover effects of the protests onto nearby jurisdictions, which may attenuate the results.

This appendix addresses both of the above concerns by aggregating from census places to counties. This aggregation alleviates concern regarding local spillover effects by extending treatment status to nearby jurisdictions. If spillover effects were to have occurred nationally, then their effect would be absorbed by the time fixed effects and, thus, would not be a concern. This aggregation allows for contiguous county estimates. If border counties tend to be similar, then this may also alleviate any concerns regarding invalid control units.

The census place level counts of police homicides and BLM protests are aggregated to counties using a 2010 census place to 2010 county crosswalk from the [Missouri Census Data Center](#) created on 4/26/2021. When a census place spans multiple counties, the county with the largest share of the population is chosen.

The estimator used here is a stacked DID design with two-way fixed effects that slightly departs from baseline model slightly:

$$\frac{Y_{e,i,t}}{N_{e,i,t}} = \mu + \sum_{k=-4}^4 \beta_k D_{k,e,i,t} + X'_{c,i,t} \kappa + \alpha_{e,i} + \delta_{e,t} + \epsilon_{e,i,t} \quad (\text{B.1})$$

where  $Y$  is the count of lethal use-of-force and  $N$  is the normalization variable (none or population) in county  $i$  during time  $t$  (yearly) within event  $e$ ,  $D_k$  takes value one during event-year  $k$  for places that have protests and zero otherwise, and  $X$  flexibly controls for population by interacting population linearly with event-population decile. The dataset is a ‘stack’ of events (not cohorts, as was used in the baseline model). To be clear, each event includes one treated county and all control places. Control counties include all counties that did not have a BLM protest during the entire sample. Meaning, each treated county will be included in only one event, but each control county will be included in every event. The time variable for each event is centered at the quarter of the first protest for both treated and control units before it is collapsed into annual sums. So  $t = 0$  always corresponds to the first protest for the treated county. Stacking by event requires event-county fixed effects  $\alpha_{e,i}$  and event-year fixed effects  $\delta_{e,t}$ . The standard errors are clustered by county since this is the level protests are assigned. The standard errors account for possible correlation within a county in the changes in lethal use-of-force.

This model departs from the baseline model given by Equation 1 when stacking by event rather than cohort and from collapsing homicides and protests into annual sums by event-year instead of using quarterly event-time. I chose to stack by event rather than cohort because it allows the contiguous county estimator to be viewed as an alternative weighting scheme.

Because there are 295 events rather than 22 cohorts, this creates an enormous computational burden. I collapse the data into event-years and omits event-years below negative five to ease the computation burden, which will slightly decrease the precision of the estimates.

Following the baseline estimates, I estimate the percentage change in police homicides per normalizing variable by dividing the estimated  $\beta_k$  from Equation B.1 by the average lethal use-of-force per normalizing variable among counties exposed to BLM protests one year prior to the first protest ( $\bar{b}_{-1}$ ). The annual percentage change in police homicides in year  $k$  is  $\frac{\beta_k - \sum_{k=-4}^{-1} \beta_k / 4}{\bar{b}_{-1}}$  and the average, annual percentage change is:

$$\% \Delta \text{Lethal Force} = \frac{\sum_{k=0}^4 \beta_k / 5 - \sum_{k=-4}^{-1} \beta_k / 4}{\bar{b}_{-1}}.$$

To gauge the robustness of the results I contrast four estimators that can be viewed as alternative weighting schemes. The first estimator is ordinary least squares:

$$w_{e,i,t} = 1 \quad \text{and} \quad N_{e,i,t} = 1.$$

The second estimator is per capita population weighted least squares:

$$w_{e,i,t} = \sqrt{\text{Population}_{c,i,t}} \quad \text{and} \quad N_{e,i,t} = \text{Population}_{e,i,t}.$$

The third estimator is contiguous counties:

$$w_{e,i,t} = 1\{Contiguous_{e,i} = 1\} \quad \text{and} \quad N_{e,i,t} = 1$$

where  $1\{Contiguous_{e,i} = 1\}$  takes value one for the treated unit in event  $e$  and all their contiguous control counties and zero otherwise. The weight is also set to zero for the entire event if the treated unit does not have any contiguous control counties. The fourth estimator uses synthetic unit weights:

$$w_{e,i,t} = \hat{\omega}_{e,i} \quad \text{and} \quad N_{e,i,t} = 1.$$

The weights are selected by event to balance lethal use-of-force trends between the treated county and the control counties in the event. Let  $i = 1$  denote the treated unit in each event, implying  $i > 1$  are control units. In particular, the weights for each event  $e$  separately minimize:

$$\operatorname{argmin} \left| Y_{e,1,t<0}^* - \sum_{i>1} w_{e,i} Y_{e,i,t<0}^* \right|$$

with constraints:

$$0.0001 < w_{e,i} < 1$$

$$\sum_{i>1} w_{e,i} = 1$$

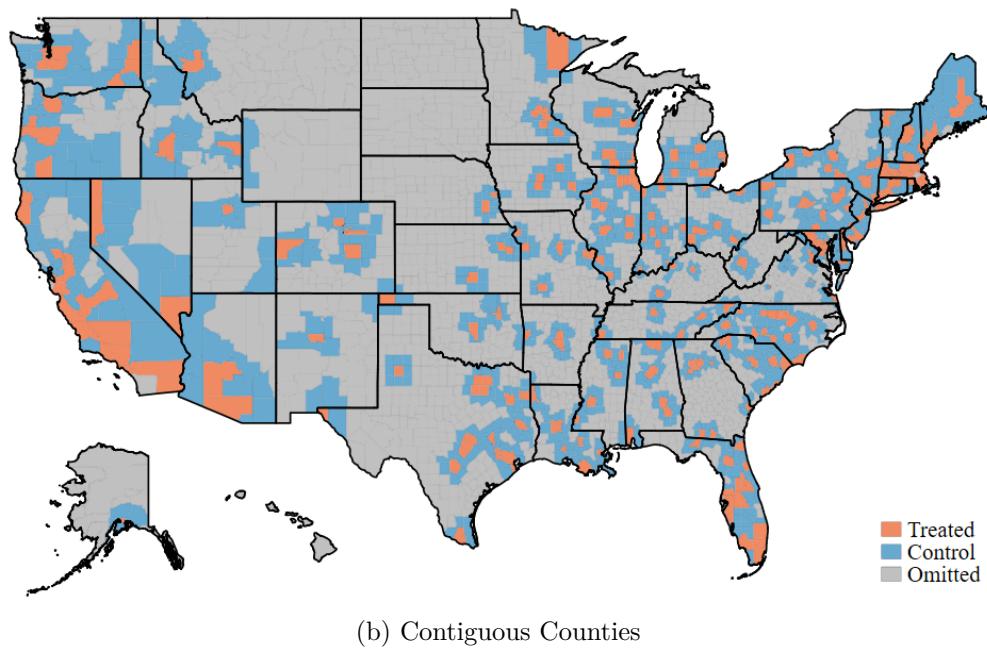
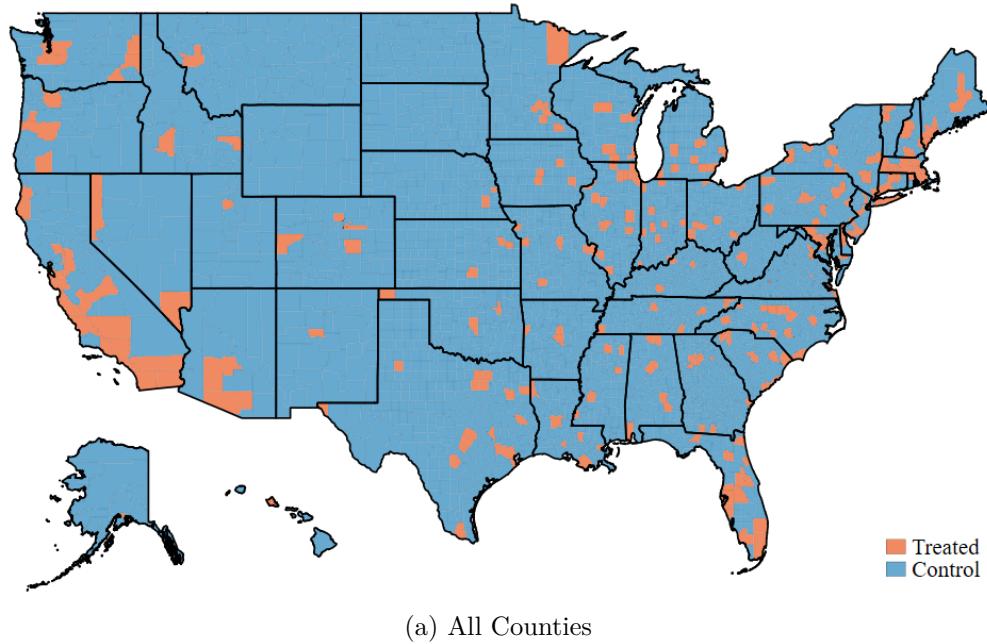
where  $Y^*$  denotes police homicides demeaned by event-county, event-time, and event-population decile. The asymptotic properties of this estimator are strongest when many units have nonzero weight (Arkhangelsky et al., 2019). A ridge penalty is conservatively set to

the stacked sample size to help evenly distribute the weight; as the ridge penalty becomes arbitrarily large, this estimator converges to ordinary least squares because equal weight is put on all counties. The synthetic unit weights are selected with the above condition rather than using inverse probability weighting (as before) because of convergence issues with logistic regressions when there is only one treated unit.

Figure B.1 maps the county level dispersion of protest status. Figure B.2 shows that the synthetic unit weights tend to emphasize contiguous counties, albeit fewer of them, which bolsters confidence in the synthetic difference-in-differences estimator.

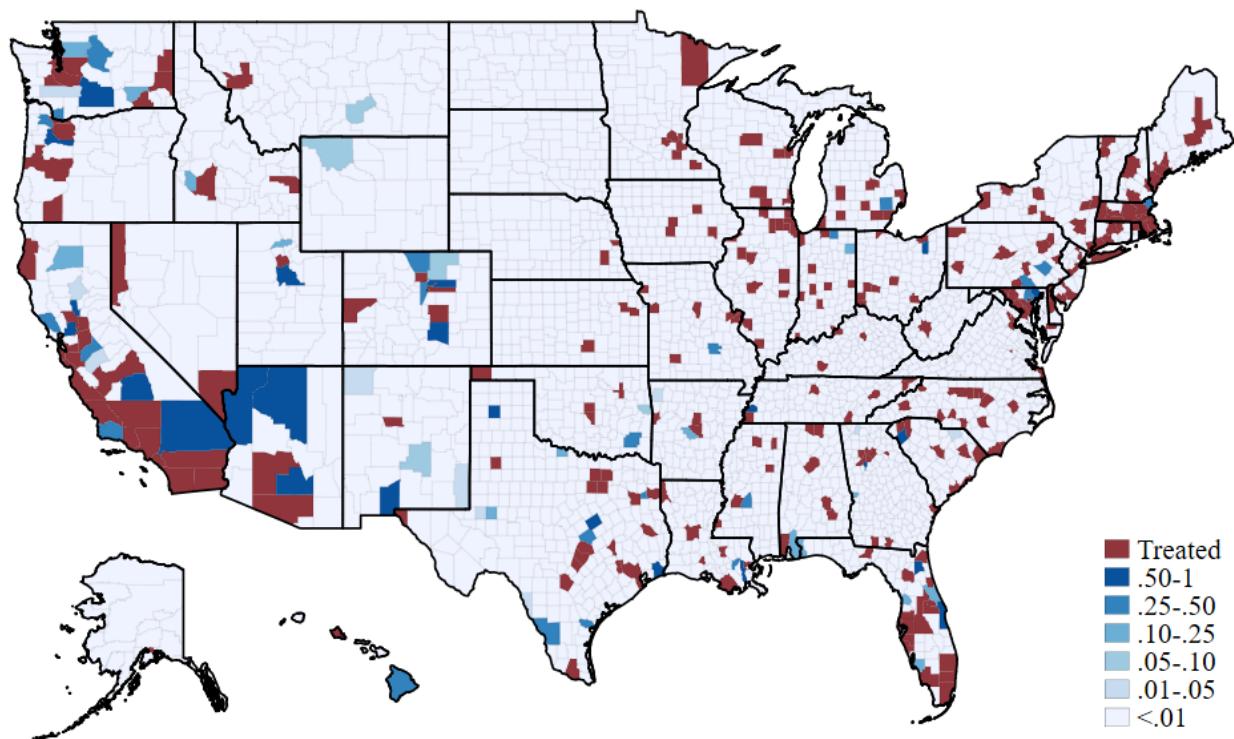
The primary findings for the county level analysis are given in Figure B.3. The principal result holds (a reduction in lethal force) at the county level using ordinary least squares, per capita population-weighted least squares, contiguous counties, or synthetic difference-in-differences. There is also a non-negligible increase in the immediate effect of the protests. Protests, therefore, may impact the behavior of both local and nearby police officers, attenuating shorter-run estimates.

Figure B.1: Maps Black Lives Matter Protests Treatment Status by County



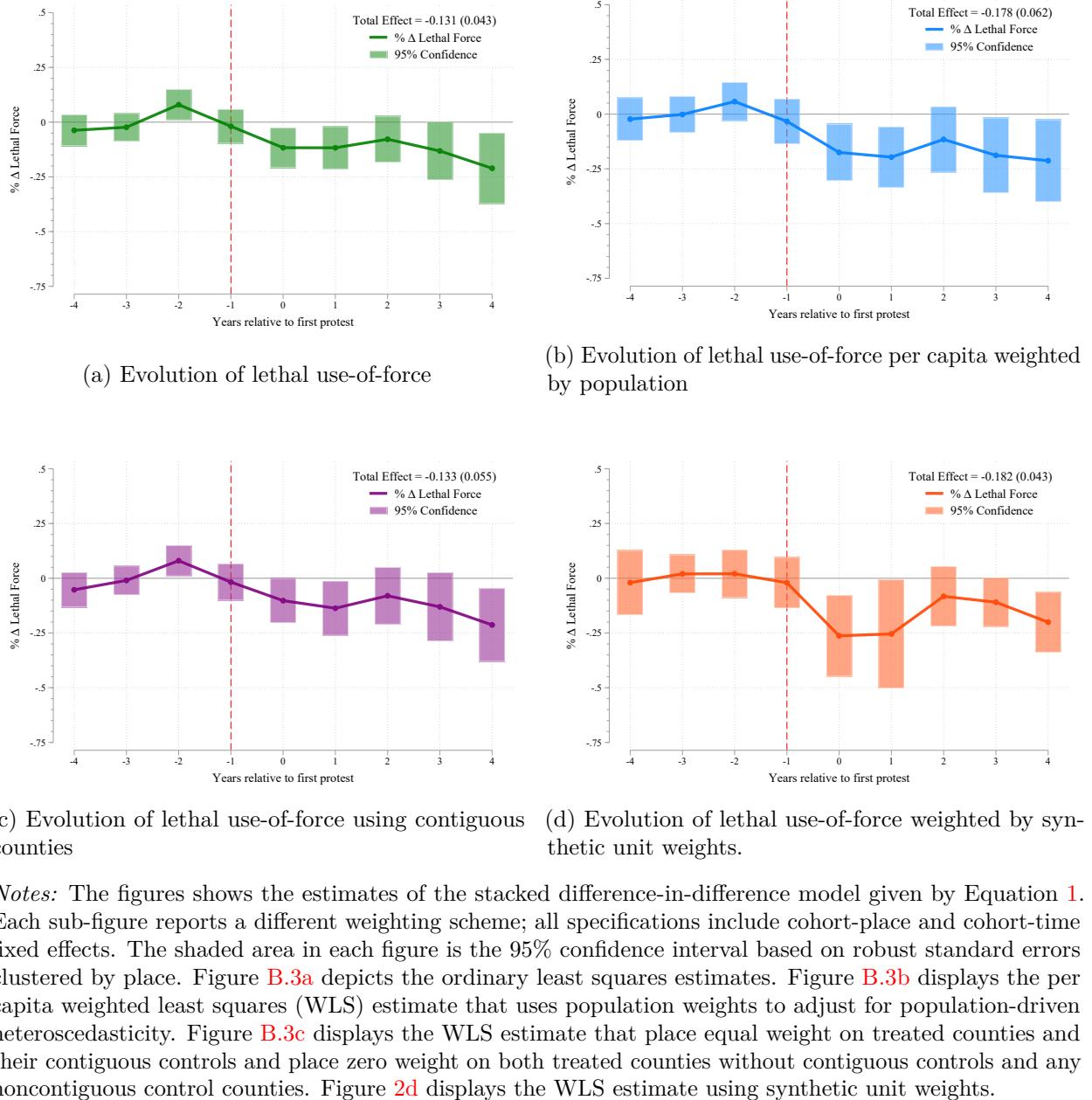
*Notes:* The figures depict treatment status by county. A county is considered treated if at least one Black Lives Matter protest is observed during the sample, 2014-2019. A county is a control if no protests are observed during the sample. Figure B.1a includes all US counties. Figure B.1b includes only treated counties and their contiguous control counties. Treated counties without at least one contiguous control county are also omitted.

Figure B.2: Distribution of Synthetic Unit Weights over Counties



*Notes:* The figure depicts the maximum synthetic unit weight assigned to a county. Each county's synthetic unit weight varies by event because the synthetic unit weights are assigned for each event individually and all control counties are included in each event.

Figure B.3: Evolution of Impact of Black Lives Matter Protests on Police Homicides



## C. Appendix – Case Studies of the 30 Most Prominent Police Scandals.

This Appendix explores heterogeneity in the impact of protests on police homicides from viral videos. Viral videos and protests are often concurrent; therefore, the baseline estimates may misattribute the effect of viral videos to protests. To gauge the importance of video recordings, I categorize the thirty of the most prominent police scandals since 2014 into three groups; one, incidents without a viral video; two, incidents with the viral video is concurrent to or precedes the first protest; three, incidents where the first protest precedes the viral video.

I provide falsification tests by leveraging the timing and location of both videos and protests; if videos reduce lethal use-of-force in places without protests, then I am likely overstating the impact of protests; likewise, if protests do not reduce lethal use-of-force without video, then I am less confident in any direct protest effect.

I measure the prominence of police scandals using the data on the subject of BLM protests as reported in the Elephrame data. The 30 case studies are chosen because they have the most protests specifically about either the name of the victim(s) or the involved officer(s).

For each of these cases, I manually research the **date of the viral video**, if any.

The **date of the first protest** corresponds to the date of the first protests with the subject of either the name of the victim(s) or the involved officer(s). To be clear, BLM protests may have already occurred in the city before the **date of the first protest** with another subject. While restricting the protests by subject may create an undercount of the number of protests, the benefit of protests unrelated to the specific incident not incorrectly categorizing protests as preceding the video outweighs the cost.

Appendix C Table C.1 list the 30 case studies in order of the total number of protests of the incident–protests where the subject is the victim(s) or officer(s) name. Protests have preceded the viral videos in over half of these cases. In many of these cases, media reports suggest the videos were made public due to public pressure from protests. It thus seems reasonable to infer that the effect of the video is part of the causal channel of protests for these specific cases. Therefore, if I do not detect heterogeneity in the effect between these groups and the estimates are fairly precise, then my confidence in the baseline results would be bolstered.

The daily lethal force data is collapsed into monthly (30 days) counts centered on the day of the first protest. The event window includes three years prior to protest and four years after protest because this is the widest event-window where the majority of events are observed in every event-year.

Four of the 30 case studies are omitted. Three of these four are omitted because they occur in the same city as a previous event. The non-fatal shooting of Paul Witherspoon and Stephanie Washington is omitted because no homicides are recorded in New Haven, CT, during the event window, the city where the scandal took place. Without variation in homicides, the fixed effects omit this event.

The control group includes all cities with populations above 20,000 for the entire sample that do not have any BLM protests during the event window. All control cities without variation in police homicides during the event window are dropped.

To estimate the impact of protests on police homicides, I use the stacked synthetic difference-in-differences estimator described in Appendix B.

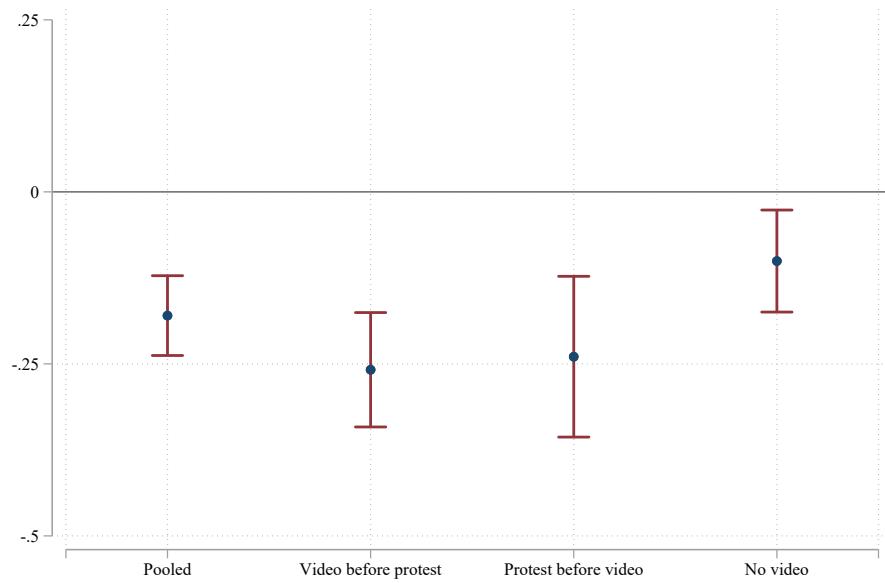
If the model is correct, BLM protests decrease lethal use-of-force regardless of video group, as shown by Figures C.1 and C.2. However, the impact of BLM protests is substantially smaller when unaccompanied by a viral video. It is unclear if this is because of the absence of a viral video, the small number of events, or a cohort effect. High-profile events without videos are more likely to occur in early cohorts (before the widespread distribution of police body-worn cameras). Only five of the thirty most prominent events do not have a video. Of these five events, three occur in 2014, one in 2015, and one in 2016. On average, events without videos occur 1.6 years prior to the other 25 events. If the later cohorts of protests are more effective, then the result may be a cohort effect, not a video effect.

Table C.1: List of High Profile Police Killings from 2014 to 2019

(1)	(2)	(3)	(4)	(5)	(6)	
#	Name of victim	Location	Total protests	First protest	Video release	Protest before video
1	Michael Brown	Ferguson, MO	133	8/9/14		
2	Philando Castile	St. Paul, MN	100	7/6/16	7/6/16	
3	Alton Sterling	Baton Rouge, LA	98	7/5/16	7/6/16	✓
4	Eric Garner	New York, NY	58	7/19/14	7/18/14	
5	Anthony Lamar Smith	St. Louis, MO	44	9/15/17	9/7/17	
6	Stephon Clark	Sacramento, CA	44	3/19/18	3/23/18	✓
7	Laquan McDonald	Chicago, IL	34	11/24/15	11/24/15	
8	Jamar Clark	Minneapolis, MN	32	11/15/15	3/30/16	✓
9	Freddie Gray	Baltimore, MD	28	4/18/15	4/21/15	✓
10	Tamir Rice	Cleveland, OH	22	11/24/14	11/26/14	✓
11	Keith Lamont Scott	Charlotte, NC	21	9/20/16	9/23/16	✓
12	Paul Witherspoon and Stephanie Washington	New Haven, CT	20	4/16/19	4/23/19	✓
13	Terence Crutcher	Tulsa, OK	19	9/19/16	9/19/16	
14	Sandra Bland	Houston, TX	18	7/17/15	5/6/19	✓
15	Sam Dubose	Cincinnati, OH	17	7/23/15	7/29/15	✓
16	Dontre Hamilton	Milwaukee, WI	14	8/25/14		
17	De'Von Bailey	Colorado Springs, CO	11	8/3/19	8/15/19	✓
18	Decynthia Clements	Elgin, IL	10	3/13/18	3/22/18	✓
19	Antwon Rose	Pittsburgh, PA	9	6/20/18	6/21/18	✓
20	Ronald Johnson III	Chicago, IL	8	12/7/15	12/7/15	
21	Atatiana Jefferson	Fort Worth, TX	8	10/13/19	10/12/19	
22	Tony Robinson	Madison, WI	7	3/6/15	5/13/15	✓
23	John Crawford III	Beavercreek, OH	7	8/30/14	9/24/14	✓
24	Mario Woods	San Francisco, CA	7	12/23/15	12/7/15	
25	Jonathan Ferrell	Charlotte, NC	7	7/19/15	8/5/15	✓
26	Tyre King	Columbus, OH	7	9/20/16		
27	David Jones	Philadelphia, PA	7	7/20/17	6/9/17	
28	Harith Augustus	Chicago, IL	6	7/14/18	7/15/18	✓
29	Ezell Ford	Los Angeles, CA	6	8/17/14		
30	Aura Rosser	Ann Arbor, MI	6	3/2/15		

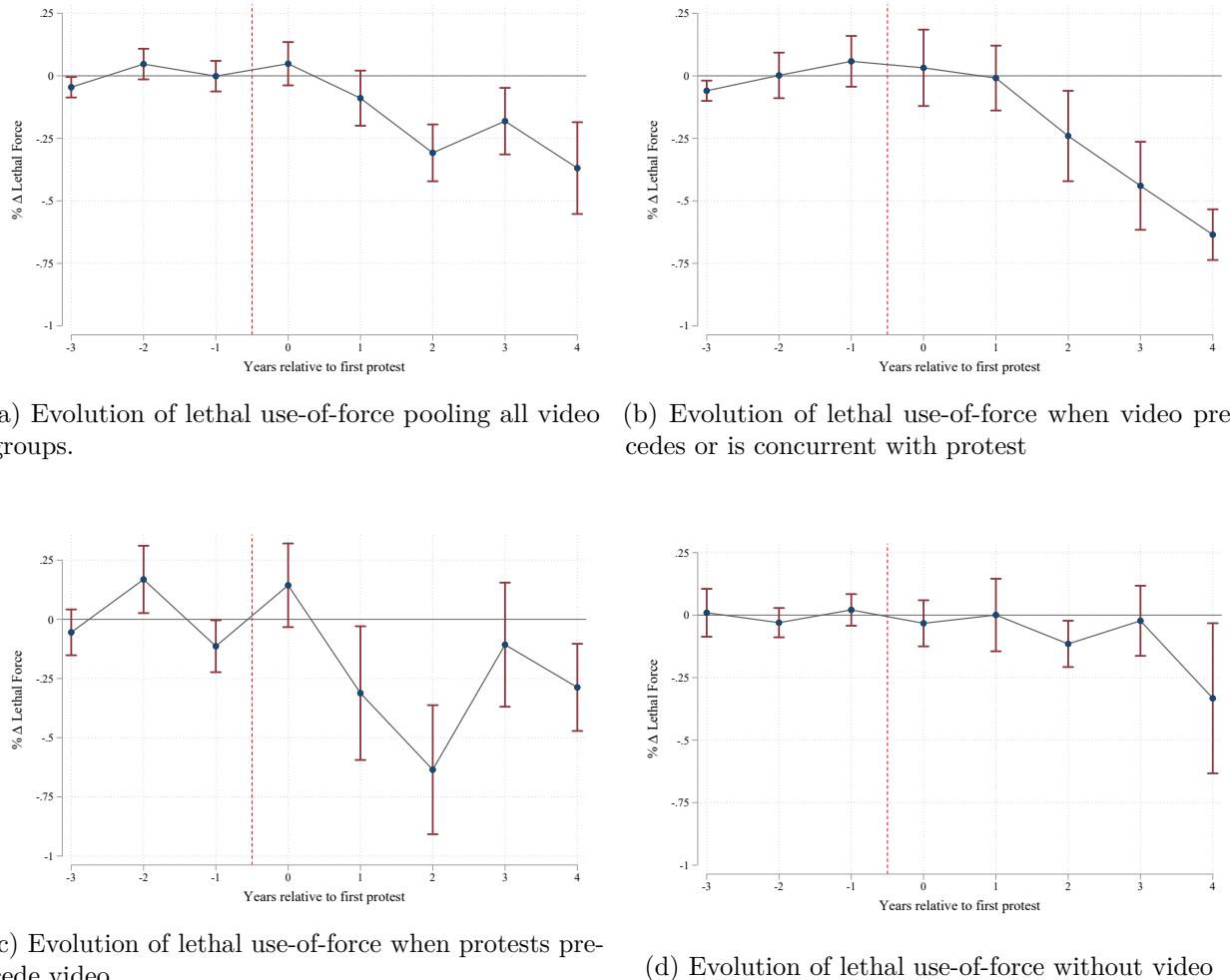
Notes: Philando Castile, Sandra Bland, and Antwon Rose were killed outside of the reported city in nearby areas below the 20,000 population screen.

Figure C.1: Heterogeneity in Black Lives Matter's Impact on Lethal Force from Video Timing Pooling Most Prominent Events



*Notes:* The figure provide estimates of the stacked difference-in-difference model given by Equation 1 by video group using only the most prominent events. All specifications include event-place, event-time, event-population decile fixed effects and are estimated with ordinary least squares. The bars are the 95% confidence intervals based on robust standard errors that are clustered by place. "Pooled" pools all video groups. "Video before protest" shows the overall estimate using only events where the video of the police homicide is made public before or on the same day as the first protest. "Protest before video" shows the overall estimate using only events where the first protest occurs before the video is released to the public. "No video" displays the overall estimate using only events where a video of the police homicide was never released to the public or does not exist.

Figure C.2: Evolution of Video Timing Heterogeneity Estimates



(a) Evolution of lethal use-of-force pooling all video groups.

(b) Evolution of lethal use-of-force when video precedes or is concurrent with protest

(c) Evolution of lethal use-of-force when protests precede video

(d) Evolution of lethal use-of-force without video

*Notes:* The figure provide estimates of the stacked difference-in-difference model given by Equation 1 by video group using only the most prominent events. All specifications include cohort-place, cohort-time, cohort-population decile fixed effects and are estimated with ordinary least squares. The bars in each figure are the 95% confidence intervals based on robust standard errors that are clustered by place. Figure C.2a pools all video groups. Figure C.2b shows estimates using only events where the video of the police homicide is made public before or on the same day as the first protest. Figure C.2c shows estimates using only events where the first protest occurs before the video is released to the public. Figure C.2d displays estimates using only events where a video of the police homicide was never released to the public or does not exist.

## D. Appendix – Event Study Distinguishing Protests from Video Recordings of Police Homicides

One may be concerned that the baseline estimates misattribute the impact of viral videos to protests. To alleviate this concern, this appendix leverage the timing and location of both videos and protests for two falsification tests; if videos reduce lethal use-of-force in places without protests, then I am likely overstating the impact of protests; likewise, if protests do not reduce lethal use-of-force without video, then I am less confident in any direct protest effect. Places with protests and videos experience significant reductions in lethal force.

The data for videos from this exercise comes from Mapping Police Violence (MPV). This data includes an indicator for whether or not there was a video recording from a video that relies heavily on the Washington Post data. Because the Washington Post data began in 2015, there is likely a severe undercount in years prior.

To accomplish the above task, I create four separate quarterly panels. For all four panels, the control group includes all census places that do not have a video recording or protest during the entire sample. All four panels are stacked by cohort as described in Section II. The distinction between each panel is from the treated group and how event time is centered.

The first panel is used to estimate the impact of protesting with a video recording on police homicides. The treated units in the first panel have both protests and videos. Event time is aligned by the quarter of the first protest by cohort.

The second panel is used to estimate the impact of protesting without a video recording on police homicides. The treated units in this panel have protests but do not have a video. Event time is aligned by the quarter of the first protest by cohort.

The third panel is used to estimate the impact of videos with protests on police homicides. The treated units in this panel have both protests and videos. Event time is aligned by the quarter of the first video by cohort.

The fourth panel is used to estimate the impact of videos without protests on police homicides. The treated units in this panel have video but do not have a protest. Event time is aligned by the quarter of the first video by cohort.

The third and fourth panels inadvertently condition on a police homicide at quarter zero because a video recording of a police homicide requires a police homicide to take place. This is an issue because conditioning on an outcome biases the difference-in-differences estimate. Therefore, I subtract one from the measure for police homicide in quarter zero for these two panels. The estimates can thus be thought of as the impact of videos on *other* police homicides.

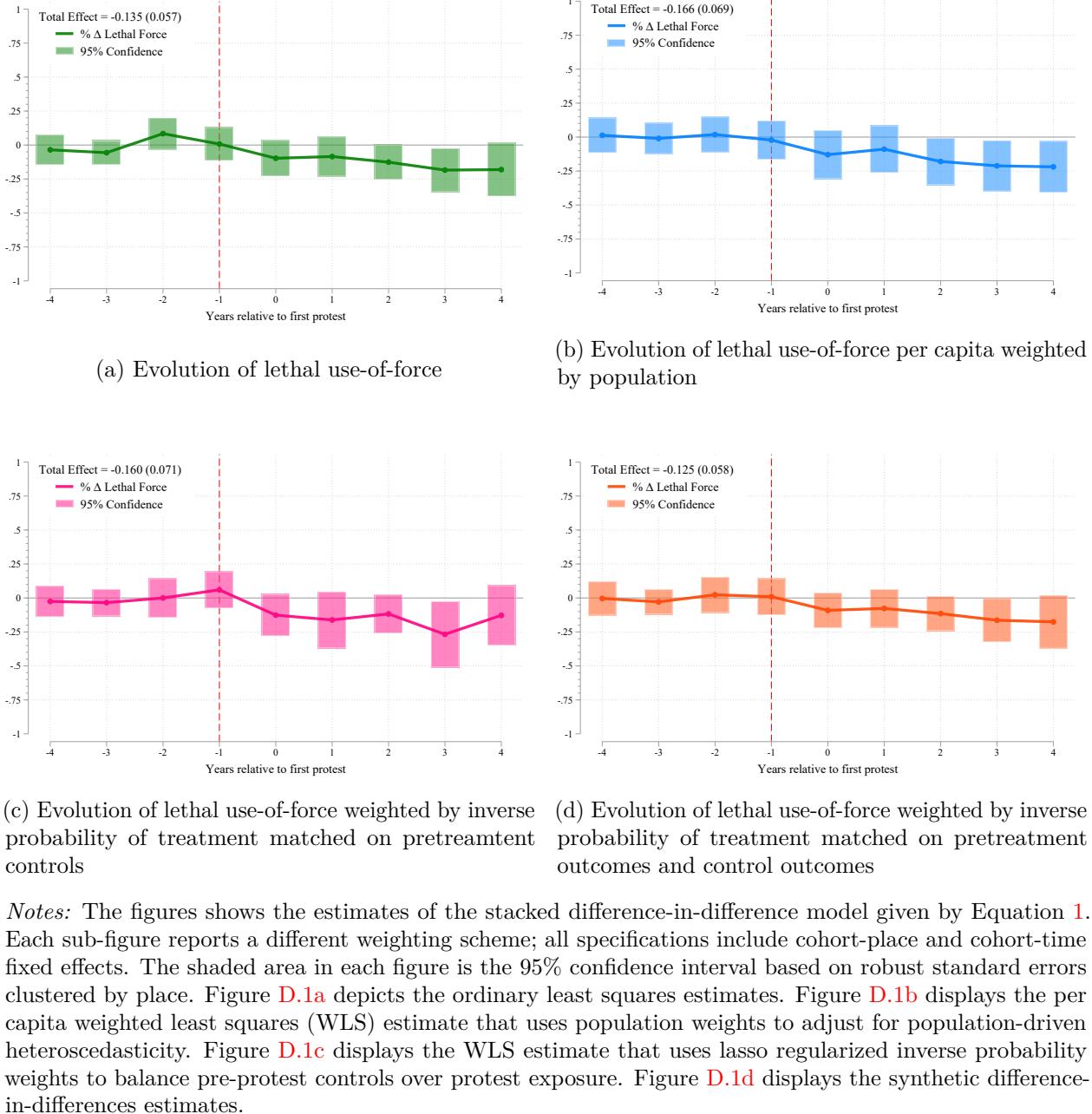
To be consistent with the baseline estimates, I report estimates using four estimators for each panel that are analogous to those described in Section II: ordinary least squares, per capita population-weighted least squares, doubly robust inverse probability weighting, and synthetic difference-in-differences. To be clear, the only difference between these sets of results and the baseline estimates is the restriction on the control group, the restriction on the treated group, and, potentially, how event time is aligned.

If the model is correct, protests decrease lethal use-of-force when accompanied by videos, which is consistent with the baseline results (see Figure D.1). Likewise, Figure D.2 shows a reduction in lethal force following video recordings of police when accompanied by protests.

The key difference between these two figures is the aligning of event-time. Notably, aligning event-time by videos for events with both protests and videos creates imprecisions and a significant negative pre-trend. One explanation could be that videos tend to follow protests for multiple reasons; one, protests were shown in Section IV.A to increase body-worn camera adoption; two, protests may pressure agencies to release body camera video to quell tension or to control the narrative; and three, protests may foster community vigilance, emboldening onlookers to record police-civilian interactions; four, there is a severe undercount of videos prior to 2015. In fact, 73 of the 94 cities with both protests and videos during the sample experienced protests at least one quarter prior to the video.

The estimates pass both falsification tests, which is consistent with protests driving the reduction in lethal force rather than viral videos in-of-themselves. Lethal force falls following protests when unaccompanied by a video when using ordinary least squares, per capita population-weighted least squares, doubly robust inverse probability weighting, or synthetic difference-in-differences (see Figure D.3). However, the reduction in lethal force is slightly smaller and much less precise than when protests are accompanied by videos. Figure D.4 shows a significant increase in police homicides following videos when unaccompanied by protests for all four specifications.

Figure D.1: Evolution of Impact of Protests with Video of Police Homicides on Police Homicides



*Notes:* The figures show the estimates of the stacked difference-in-difference model given by Equation 1. Each sub-figure reports a different weighting scheme; all specifications include cohort-place and cohort-time fixed effects. The shaded area in each figure is the 95% confidence interval based on robust standard errors clustered by place. Figure D.1a depicts the ordinary least squares estimate. Figure D.1b displays the per capita weighted least squares (WLS) estimate that uses population weights to adjust for population-driven heteroscedasticity. Figure D.1c displays the WLS estimate that uses lasso regularized inverse probability weights to balance pre-protest controls over protest exposure. Figure D.1d displays the synthetic difference-in-differences estimate.

Figure D.2: Evolution of Impact of Video of Police Homicides with Protests on Other Police Homicides

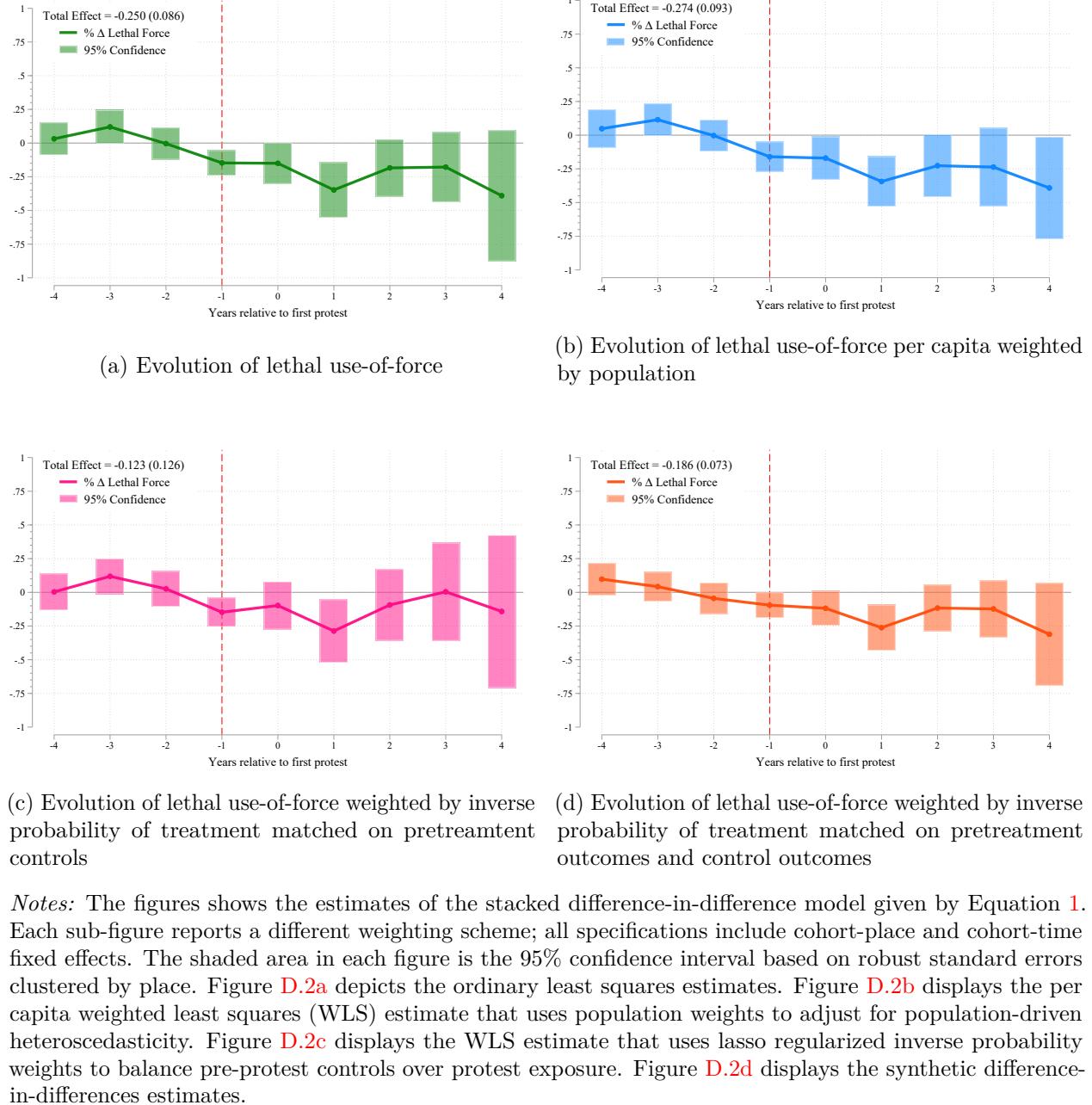
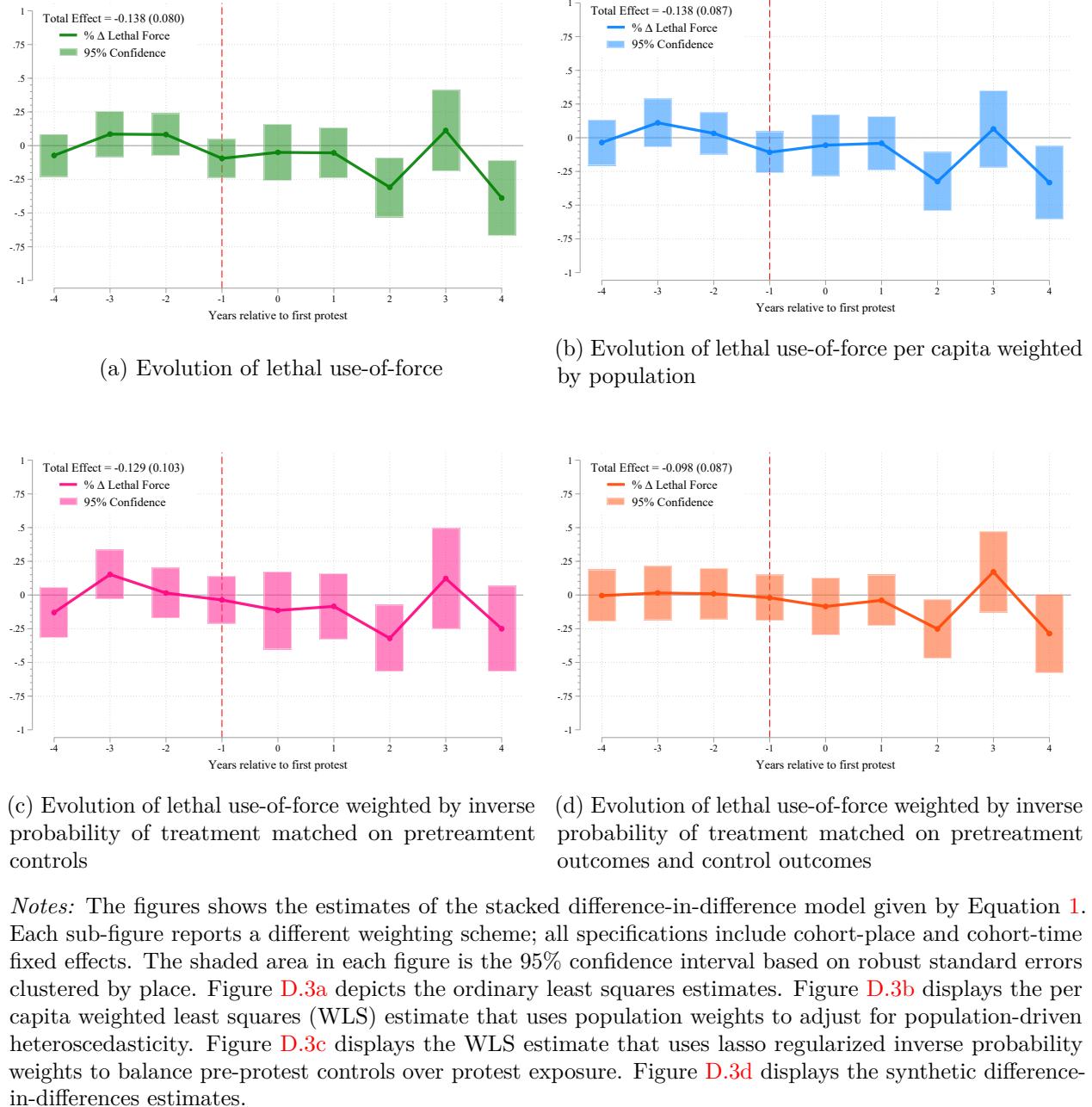
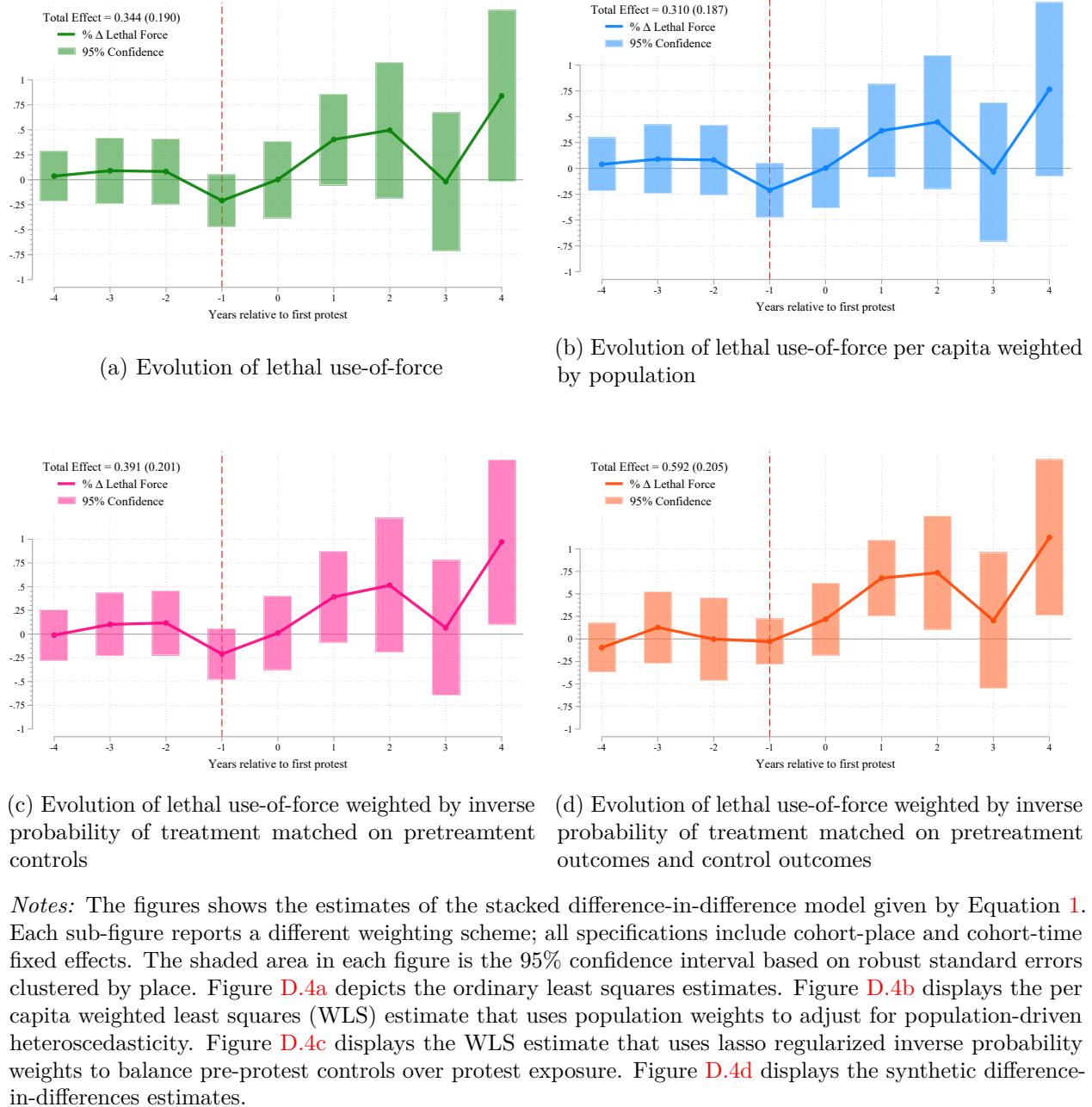


Figure D.3: Evolution of Impact of Protests without Video of Police Homicides on Police Homicides



*Notes:* The figures show the estimates of the stacked difference-in-difference model given by Equation 1. Each sub-figure reports a different weighting scheme; all specifications include cohort-place and cohort-time fixed effects. The shaded area in each figure is the 95% confidence interval based on robust standard errors clustered by place. Figure D.3a depicts the ordinary least squares estimates. Figure D.3b displays the per capita weighted least squares (WLS) estimate that uses population weights to adjust for population-driven heteroscedasticity. Figure D.3c displays the WLS estimate that uses lasso regularized inverse probability weights to balance pre-protest controls over protest exposure. Figure D.3d displays the synthetic difference-in-differences estimates.

Figure D.4: Evolution of Impact of Video of Police Homicides without Protests on Other Police Homicides



*Notes:* The figures show the estimates of the stacked difference-in-difference model given by Equation 1. Each sub-figure reports a different weighting scheme; all specifications include cohort-place and cohort-time fixed effects. The shaded area in each figure is the 95% confidence interval based on robust standard errors clustered by place. Figure D.4a depicts the ordinary least squares estimate. Figure D.4b displays the per capita weighted least squares (WLS) estimate that uses population weights to adjust for population-driven heteroscedasticity. Figure D.4c displays the WLS estimate that uses lasso regularized inverse probability weights to balance pre-protest controls over protest exposure. Figure D.4d displays the synthetic difference-in-differences estimates.