Do Cops Know Who to Stop?

Assessing Optimizing Models of Police Behavior with a Natural Experiment

David S. Abrams, Hanming Fang and Priyanka Goonetilleke*

July 15, 2021

ABSTRACT

The standard economic model of police stops implies that the contraband hit rate should rise when the number of stops falls, *ceteris paribus*. We provide empirical corroboration of such optimizing models of police behavior by examining changes in stops and frisks around two extraordinary events of 2020 - the pandemic onset and the nationwide protests following the killing of George Floyd. We find that hit rates from pedestrian and vehicle stops generally rose as stops and frisks fell dramatically. Using detailed data, we are able to rule out a number of alternative explanations, including changes in street population, crime, police allocation, and policing intensity. We find mixed evidence about the changes in racial disparities, and evidence that police stops do not decrease crime, at least in the short run. The results are robust to a number of different specifications. Our findings provide quantitative estimates that can contribute to the important goals of improving and reforming policing.

* Author affiliations and contact information. Abrams: University of Pennsylvania Law School and Wharton (dabrams@upenn.edu) Fang: University of Pennsylvania Department of Economics (hanming.fang@econ.upenn.edu) Goonetilleke: University of Pennsylvania Department of Economics (goonep@sas.upenn.edu)

I. Introduction

Gary Becker (Becker, 1968) deserves much of the credit for broadening economics to include the study of crime and race and suggesting the use of outcome tests. But Knowles, Persico, & Todd (Knowles, Persico, & Todd, 2001, heretofore KPT), in an influential study, first suggested the use of what is called a "hit rate" test to distinguish racial prejudice from statistical discrimination as the underlying cause for racial disparities in highway searches by police officers. An implication of KPT and almost all subsequent models on the subject is of diminishing marginal returns to stops (see Section II for more details). That is, if stops decrease substantially, the likelihood of contraband discovery should rise, ceteris paribus.

This is the first paper with the opportunity to test this assumption empirically, as exogenous changes to policing are rare. But 2020 was a rare year, and brought not one, but two events where police stops dropped tremendously and rapidly. The first event was the COVID-19 pandemic onset in March 2020, and the second was the nationwide protests and looting that followed the killing of George Floyd by a police officer on May 25, 2020 in Minneapolis.

Using extremely granular data from Chicago and Philadelphia, we examine these events, relying mostly on the latter, and find that hit rates generally increased as police stops plunged, according to the predictions of KPT. At the same time, the rate of legally unfounded stops fell. This fact along with other evidence we discuss below suggests that the impact wasn't due to a change in police effort, nor composition of individuals on the street, changes in police deployment, or crime. While the total number of police making stops and stops per officer both fell, the latter seems to be more responsible for the increased hit rate. We examine changes in hit rate by race, with ambiguous results. No evidence for a deterrent impact of police stops and frisks is found.

KPT kicked off a large literature in economics studying racial bias in all aspects of law enforcement and the judicial system including policing, highway patrols, bail setting, sentencing, probation, parole, among others see e.g. (Abrams, Bertrand, & Mullainathan, 2012; Anwar & Fang, 2006; Ayres, 2002; Bjerk, 2007; Devi & Fryer Jr., 2020; Dharmapala & Ross, 2004; Durlauf, 2006; Heaton, 2010; Knox & Mummolo, 2020; Persico, 2002;

Ridgeway, 2006; Sanga, 2009). Now 20 years old, the "hit rate" analysis proposed in KPT has been used to evaluate the race-neutrality of policing in an array of cities, and has elicited a great deal of theoretical and empirical response. Although many aspects of KPT have been challenged, almost all subsequent work models officers as agents seeking to maximize a police objective function related to stops or searches. The objectives of the police typically include a weighted average of a legally justified part of the payoff, e.g. finding contraband, and an illegitimate part of the payoff, e.g. racial prejudice.

We use the exogenous changes in policing due to the pandemic and protests to investigate the predictions of policing models. In addition to testing the models, our estimates are important parameters to know when making policy decisions about policing. While the change in police stops is large for both events, frisks are almost flat during the pandemic onset, and mobility declines sharply at that time, providing potential confounds. Thus most of the analysis is focused on the protests, when mobility changes are small but frisks plummet alongside stops.

We perform several additional tests of the protest period in order to rule out the possibility that our results are due to simultaneous changes. In order to test possible changes in traffic composition, we obtain hospital data including the age distribution of individuals involved in accidents, and find no significant change. Even though the numbers and some measure of the composition of the relevant population is static it is still possible that the share of potential criminals grows. Examining crime data, we find that it is mostly flat or decreasing in this period, depending on the category.

Changes in policing could also account for the increase in hit rates. This includes a change in police deployment and effort per officer following the changes. By focusing on a constant group of officers we rule out changes in deployment and find that the increased hit rate is largely due to the decline on the intensive rather than extensive margin.

Although stop-and-frisk policing has long been justified based on an asserted impact on crime, we find no evidence for that in our data. Lagged crime actually falls after the large drop in police stops and frisks in Philadelphia. While not the focus of the paper, we examine racial disparities and find mixed results over the protest period. We perform a number of robustness checks - with varying time windows, using stop hit rates (in addition to frisk hit rates), and perform simple before-after comparisons. All of the results are consistent with the main findings and in some specifications more robust.

In addition to testing policing models, these results should help with the crucial task of policy decisions about policing. Understanding the efficacy and sensitivity of police stops to abrupt changes are needed to evaluate their utility.

The rest of the paper proceeds as follows. Section II provides a brief background on police stops. Section III introduces the data on police stops in Chicago and Philadelphia. Section IV includes the main analysis and Section V adds robustness checks with a discussion of potential threats to validity. Section VI concludes.

II. Background

Becker's outcome test relies on the following idea: if a police officer is racially prejudiced against a minority group, he/she would then choose to stop and/or search a minority group member with less convincing indications for the presence of contraband. Thus, in the aggregate, the contraband finding rates among the stopped/searched minority group members should be lower. Therefore, the comparisons of the ex post outcomes against different groups can be suggestive of the presence of racial prejudice by the police. However, as argued in the literature following KPT, the infra-marginality problem may complicate the applications of the outcome test. To understand the infra-marginality problem, recall that conceptually Becker's outcome test is based on the idea that a police officer will search a driver if the suspicion level is above a threshold; an officer prejudiced against minority drivers will use a lower suspicion threshold for minorities than for white drivers. Thus, theoretically the contraband finding rate among the marginal minority drivers and that among the marginal white drivers are indicative of the officer's prejudice. However, the outcome tests often compare the average contraband finding rates against different groups. The comparisons of group averages may not be in the same direction as the comparisons of the marginals. This is the issue of the inframarginality problem.

KPT resolves the infra-marginality problem by presenting an equilibrium "matching pennies" game between police officers and drivers. The equilibrium of the "matching pennies" game is in mixed strategies, and minority drivers will carry contraband with lower probability in this mixed strategy equilibrium if and only if the police officers are prejudiced against them. Moreover, in this mixed strategy equilibrium police officers will search the minority drivers with higher probability, which is the reason driving down the probability of carrying contraband by the minority drivers. A feature of the KPT model, however, is that the marginal drivers and the average drivers are the same because in the mixed strategy equilibrium all drivers of the same race are carrying contraband at the same rate, even if the drivers differ in their propensity to commit crimes.

Subsequently, Anwar and Fang (2006) presents a model of policing behavior in which officers decide whether to search a driver after observing signals about whether the drivers may be carrying contraband. To the extent that the police observe more or less suspicious signals regarding the drivers' potential guilt, the officers who are interested in maximizing, at least as part of its objective function, the contraband finding rate, will be searching the drivers only if their suspicion for the driver carrying contraband exceeds a threshold. That is, they will be allocating their search effort only on those that they deemed to be more suspicious of carrying contraband. If officers are prejudiced against certain groups of drivers, then they will be using a lower suspicion threshold against drivers from that group. Since drivers within the same racial group are heterogeneous in their level of suspicion, the infra-marginality problem exists in this setting. Anwar and Fang (2006) address this issue by introducing officers of different races and use offices of a given race as a benchmarking to assess the relative prejudice of one group of officers against another group of officers.

While the mechanisms underlying the KPT model and the Anwar and Fang (2006) model differ substantially, they share the following key feature and prediction. First, both models assume that the police officers are rational and are trying to optimize an objective that includes contraband finding as one of its components. Second, both models would predict that, if for some exogenous reasons the costs of stopping or searching drivers (or pedestrians) were to go up, the contraband finding rates against all drivers should go up. In KPT model, this prediction emerges through the endogenous response of the drivers who are deciding whether to carry contraband. As the officers' cost of searching vehicles increase, it is necessary for the drivers to increase their probability of carrying contraband to ensure that the officers are indifferent between searching and not searching in the mixed strategy equilibrium. In Anwar and Fang (2006), an increase in search cost by the officers will make the officers increase the suspicion threshold of the drivers that they search. This increase in the marginal suspicion threshold will necessarily increase the average search success rate of each group of drivers.

Given the prominence of the rational choice framework in the study of police behavior, it is of great value to examine the empirical foundation of the assumption that police indeed aim to maximize at least to some extent the contraband finding rate in deciding whom to stop and search.

The focus in this paper is policing in Philadelphia, which in many ways is typical of large American cities over the last decade. Stop-and-Frisk policing is a tactic that has been widely employed and is viewed by many in law enforcement as decreasing crime, although the authors are aware of no well-identified study to this effect. As the name implies, officers stop individuals whom they suspect of potential wrongdoing. If warranted, they may then proceed to frisk and potentially arrest the suspect.

Legally, these stops of individuals are known as Terry stops, after the 1968 Supreme Court case *Terry v. Ohio*. The case established that officers may stop individuals if they have reasonable suspicion of involvement in criminal activity. A frisk is allowed if there is reasonable suspicion that the individual may be armed. In the case of a vehicle stop, while the same reasonable suspicion requirement applies for a frisk of the driver and passengers, officers are allowed to conduct a search of the vehicle if there is probable cause to believe that evidence of any criminality, including a drug violation, is concealed within the vehicle. The Department of Justice, the ACLU and other organizations have investigated or sued dozens of police departments since 2000 for unlawful, excessive, and racially disparate use of stop and frisk. Much of the recent economics literature on police stops has informed these cases, including *Bailey v City of Philadelphia*, which led to a settlement agreement that is still in force in 2021.

A. Related Literature

Dharmapala and Ross (2004), Anwar and Fang (2006), and (Antonovics & Knight, 2009) discussed the possible shortcomings of the KPT model. Dharmapala and Ross (2004) point out that KPT's test does not generalize if potential drug carriers may not be *observed* by the police or if there are different levels of drug offense severity. Under those circumstances KPT's test fails because the infra-marginality and omitted variables problems re-emerge. More specifically, the equilibrium of the KPT model under those circumstances may involve a group of motorists carrying drugs with probability one (being a ``dealer'') even when they are searched with probability one whenever the troopers observe them. If the probability of being a ``dealer'' is higher for minorities, then the average success rate against minorities should be greater than that for whites under statistical discrimination, and equal average success rates would actually indicate taste discrimination, contrary to KPT's conclusion.

Antonovics and Knight (2007) argued that KPT's test may not be robust when its model is generalized to allow for trooper heterogeneity. As in Anwar and Fang (2006), they also proposed using data with both motorist and officer information, and they run a Probit regression using data from the Boston Police Department where the dependent variable is an indicator for whether a search took place for a given stop, and the explanatory variables include some observable characteristics of the driver and officer and a dummy variable indicating whether there is a racial mismatch between the officer and the driver. In their baseline regression, they find a positive coefficient on the ``racial mismatch'' variable, indicating that officers are more likely to conduct a search against motorists of races different from their own. They interpret this finding as evidence of racial prejudice. However, as Anwar and Fang (2006) argued, their interpretation of the evidence may be misleading as it is possible to rationalize their findings simply by differences in the signal distributions officers receive about the motorists.

Two recent papers have sought to empirically test the underlying assumption that police officers face diminishing returns to search. Feigenberg and Miller (2021) estimate a between officer Search Productivity Curve (SPC) to determine whether there is an equity-efficiency trade-off using data on traffic stops for speeding violations conducted by Texas Highway Patrol troopers. They find that the relationship between the search rate and unconditional hit rate (hit rate as a proportion of stops) is roughly linear ie that the conditional hit rate is roughly constant across troopers with different search rates. Meanwhile, Gelbach (2021) calculates the relationship between officer-level search rates and conditional hit rate for Florida and Harris County Texas. He finds mixed results. In Florida the relationship is negative, consistent with officers facing diminishing returns to search. Meanwhile in Harris County the relationship is positive for white and Hispanic drivers which is taken as suggestive evidence that the Becker framework may not be appropriate in this setting.

However, under the optimizing models used in the literature diminishing returns applies at the officer level. Due to the rarity of exogenous shifts to policing, these papers use between officer variation as a proxy for within-officer changes. As Feigenberg and Miller (2021) note, in the absence of strict monotonicity, the between-officer SPC may not coincide with the within-officer SPC if for example search rates were correlated with the ability of an officer to identify suspects.

Besides the ``outcome test'' approach, a large field of literature has used a different statistical test, known as the ``benchmarking test,'' to test whether officers impose disparate treatment on motorists of different races. The benchmarking test typically compares the shares of racial or ethnic minorities in the population to their shares in the sample of motorists selected for discretionary stops and searches by police. The main drawback of the benchmark test is that it cannot determine if racial disparities arise out of racial prejudice or statistical discrimination. Furthermore, the benchmark test suffers from two main problems. The first problem is called the *denominator* problem, which refers to the question of what should be the right benchmark to compare the stop and search rates. It ideally should be the racial or ethnic composition of drivers on the road, but such information is typically unavailable. The second problem is the *omitted-variables* problem. If there exist certain characteristics whose distributions are correlated with motorists' race or ethnicity and if such characteristics may be observed by police but not available to researchers, benchmarking tests will not be completely informative about whether motorists' race affected the search decision.

III. Data

One of our two main data sets comes from the Philadelphia police department and contains suspect-stop level data[1] about all pedestrian police stops within the city from January 1, 2015 - December 31, 2020. After we remove the 2 observations for which location data is incomplete and the 499,638 observations for which race is missing or unknown, the data set comprises 2,321,622 stops. The data is rich and includes information on timing, location, ID of the police officer conducting the stop, demographic information on the individual stopped, whether a frisk, search or arrest occurred, vehicle information, any contraband discovered, and a narrative field. We have similar data for Chicago, spanning from January 1 2016 – December 31 2020 and comprises 533,471 stops.

In 2020, the onset of the COVID-19 pandemic dramatically impacted almost all aspects of life, including policing. This is reflected in the top panels of Figures 1a and 1b, which display an index of police stops for 2020 (solid black) and 2015 - 2019 (dashed gray) for Chicago and Philadelphia, respectively. The figures are normalized so that the stop level from January 1-7 of each series equals 100. The red vertical line indicates the start of the week in which the Philadelphia Police Department announced a new policy for nonviolent incidents while the green line marks the start of the decline in police stops following the protests and looting in response to the George Floyd killing. There was a major decline in pedestrian police stops in Chicago of 61%in mid-March and in Philadelphia of 34%. After increase back towards normal levels, there was again a dramatic decline in stops of 44% in Chicago and 74% in Philadelphia following the killing of George Floyd.[2] The policing changes around these two unanticipated events are the focus of the paper.

The bottom panels of Figures 1a and 1b show analogous changes in pedestrians frisks. In Chicago the drop in frisks was large (49%) and long-lasting after the pandemic onset, only returning to similar levels just before the Floyd killing. Frisks then dropped sharply, by 39%. In Philadelphia, the pattern for frisks is somewhat different for the first event, with a

---

[1] Each stop may involve multiple suspects and individual suspects may be stopped multiple times, so the finest-grained level of analysis is at the suspect-stop level.
[2] We refer to the dramatic decline in police stops surrounding the pandemic and the protests as the two "events" in this study.

sharp decline followed swiftly by a return to prior levels of frisks. But the period around the protests is very similar to stops with a precipitous decline of 74%, and a new lower level for frisks that stays relatively constant through most of the summer.

Figure 2 provides the analogous information for stops and searches of vehicles, rather than pedestrians. The patterns are similar. In Chicago vehicle stops fell by 63% in mid-March and 48% at the beginning of June. In Philadelphia the falls were 50% and 80% respectively. The fall in vehicle searches is even greater than pedestrian frisks. Following the pandemic, searches in Chicago fell by 59% while in Philadelphia they dropped by 31%. The decline following the onset of the protests was 48% in Chicago and 82% in Philadelphia. Appendix Table A1 reports results from regressions of log stop and frisk totals using both difference-in-difference (columns 1-4) and single difference specifications holding the set of officers constant (columns 5-8). The figures establish the sudden changes in policing during these two moments in 2020, but the key to learning from them will be understanding whether there were other relevant changes at the same time.

Table 1 begins to shed light on that by providing summary statistics for pedestrian and vehicle stops around each of the two events. For both Chicago and Philadelphia, the pandemic is defined as starting on March 16, 2020. “Before” is the 6 weeks before this date and “After” is the 4 weeks after. The George Floyd Protests are defined as beginning on May 29, 2020 in Chicago and June 3, 2020 in Philadelphia. For both cities “Before” is the 3 weeks prior to the city specific start date and “After” is the 6 weeks following.[3] While the number of stops declines sharply after both events, it is worth examining changes in their characteristics separately for each event.

For the Pandemic event, the share of stops leading to a frisk increases substantially in both cities. For pedestrian stops the frisk rate increase by about 7.5 and 9 percentage points in Chicago and Philadelphia respectively. For vehicle stops the increase in the search rate was around 4.5 percentage points in both cities. The Philadelphia Police Department (PPD) announced on March 17, 2020 a change in policy with respect to non-violent incidents in response to COVID-19 (Melamed & Newall, 2020). In order to reduce the number of people

[3] The “Before” and “After” periods vary slightly for the two events and are chosen to balance maximizing the data available, while focusing on the events and not other changes.

taken to Police District or kept in prison due to inability to pay bail, the PPD would no longer arrest individuals for certain non-violent offenses but would instead swear out a warrant to be used as the basis for an arrest at a later time. While the Chicago Police Department did not officially announce a policy change, media sources reported that a similar unofficial change in policy appears to have been enacted. Besides the decision to deemphasize certain low-level offenses, there was also likely greater caution taken on the part of police officers at this time, for fear of contracting the coronavirus. This likely played a role in both the reduction in stops as well as the increasing share that led to frisks.

While the Black share was relatively unchanged in Chicago around the pandemic onset, in Philadelphia the Black share increased by around 10% for both pedestrian and vehicle stops. Mean age of vehicle detainees dropped by more than 1.5 year in both cities around this event.

We use two different measures of contraband. "All contraband" includes drugs, weapons, and stolen property. Drugs make up the bulk of this category. "Guns only" just includes firearms. In Chicago, hit rates generally increased with particularly large increases in hit rates for racial groups other than Black. Contraband discovery from vehicle searches fell very slightly (0.2 percentage points) driven by a drop in the hit rate for Black drivers of 1.6 percentage points. Meanwhile, in Philadelphia the only overall hit rate that increased at the pandemic onset was that of guns discovered in vehicle searches, the others all declined. In Philadelphia's pedestrian stops the difference across racial groups was particularly stark with proportionally large decreases in hit rates for frisks of Black pedestrians (2.9 percentage points for contraband and 0.9 percentage points for guns) while the hit rate from non-Black pedestrians increased (1.3 percentage points for contraband and 1.8 percentage points for guns).

Turning to the policing change around the protests. In Chicago the change in the Black share for pedestrian stops was against quite small, however there was a 10% increase in the Black share of vehicle stops. In comparison, in Philadelphia the change in Black share is somewhat smaller in magnitude than the earlier event, and also in the opposite direction - decreasing after the protests. The mean age was relatively unchanged in Chicago but increased in Philadelphia by almost 3 years for pedestrian stops. Since each of these measures is a function of both individuals available to stop as well as police decisions on whom to stop, one cannot make any inference from these comparisons alone. In Section IV we examine additional evidence of the source of these changes.

Pedestrian frisk rates and vehicle search rates are also more stable around the protests with a slight decline in each around the time of the protests. Gun hit rates in both cities rise appreciably for pedestrian stops and in Philadelphia for vehicle stops as well. For all contraband hit rates, other than a slight fall in Philadelphia for pedestrian frisks (1.2 percentage points), these also generally rise. This drop in the contraband hit rate in Philadelphia was driven by a large fall of 5 percentage points in the hit rate for non-Black pedestrians

Taken together we see larger changes in detainee characteristics around the pandemic event, even though the magnitude of the change in stops was far greater around the protests. This suggests that it may be more fruitful to focus on the protests, but in Section IV we will examine evidence from both events.

For analysis related to the pool of potential suspects from which the police choose their stops we make use of two additional type of data. The first is crash data. In Chicago we obtain this from the Chicago Police Department. This data set includes both crashes directly recorded at the scene by the responding police officer and those which are self-reported by the driver(s) involved. This dataset includes time and location of the crash as well as demographic information of the individuals involved. For Philadelphia we obtain this from the Pennsylvania Trauma Outcomes Study (PTOS), which is Pennsylvania's central trauma registry. PTOS includes patients treated in a Pennsylvania trauma center who meet the inclusion criteria: admission to the intensive care unit or step-down unit, hospital length of stay longer than 48 hours, hospital admissions transferred from another hospital, death on hospital arrival or during admission, and transfers out to an accredited trauma center. The data includes information on cause and severity of injury, demographic information of the patient and the county of the trauma center.

Additionally, we use data on individual movement (mobility data) from Google Community Mobility Reports. This data is generated from the mobile devices of individuals who have turned on Location History for their Google Account. The daily data is aggregated to the county level and reported relative to a baseline period, adjusted for day of week. It is available for 6 different location types: Grocery and pharmacy, parks, transit stations, retail and recreation, residential and workplaces.

IV. Main Results

Our first analysis is straightforward. We estimate the impact of a large change in police frisks on the share of frisks that result in contraband discovery (the "hit rate"). If officers maximize the likelihood of contraband detection in the marginal suspect, then a large decline in frisks should substantially increase the hit rate. To do so, we first perform a difference-in-difference analysis, comparing the change in hit rate around the events in 2020 to the same time period in prior years when there were no such events:

$$contraband_i = \alpha + \beta_1 after_i * treat_i + \beta_2 after_i + \sum_{k=2015}^{2020} (\gamma_k * year_{ik}) + \sum_{j=1}^{66} (\delta_j * area_{ij}) + \theta_i + \epsilon_i \qquad (1)$$

Here *i* indexes suspect-stops or suspect-frisks (it will be different for each suspect detained in a stop, and each separate time a suspect is detained), $contraband_i$ is an indicator for contraband discovery, $after_i$ is 1 if the stop occurred after the calendar date of the relevant event regardless of year, while $treat_i$ is 1 if the stop occurred in 2020. $Year_{ik}$ is a year dummy and *area* is a dummy for the police region of the city where the stop occurred (one of 70 "sectors" in Chicago and one of 66 "PSAs" in Philadelphia). θ_i is a series of demographic dummies for age, race and gender of the suspect.

Table 2 reports results from estimating equation (1) for pedestrian and vehicle frisks from 2015 to 2020 using the two separate events. As discussed in Section II, reasonable suspicion of weapons is the legal justification for a frisk. Since firearms are most of the weapons found as contraband, this is our primary outcome. Since the bulk of contraband does not include weapons we also examine this outcome as well.

We focus our analysis initially on the protests, Panel A of Table 2. Although only two of the results are statistically significant at conventional levels all but one point in the expected direction - the vast decline in frisks corresponds to an increased hit rate. The magnitude is large, relative to the low mean hit rates. For example, the 2.5 percentage point increase post protests in gun hit rate in Chicago is equivalent to 86% of the mean over this period of 2.9%, while the 1.3 percentage point increase in Philadelphia is equal to 72% of the mean.

The gun hit rate from vehicle searches in Philadelphia rises 1.7 percentage points post-protests, higher in magnitude, but smaller compared to the mean of 3.1% at 55% while the increase in Chicago is only 0.4 percentage points. The results using all contraband as the main outcome are not as large proportionally although the 3.4 percentage point increase for vehicle stops in Chicago is statistically significant at the 5% level. In Philadelphia the results for pedestrian stops is actually slightly negative.

The results from the pandemic are shown in Panel B. In Chicago the changes were generally smaller and statistically insignificant. The exception is the hit rate for guns from pedestrian stops in Chicago which had a statistically significant increase of 2.1 percentage points which is 115% of the mean for these dates. Meanwhile, in Philadelphia three of the results are negative although statistically insignificant. The only hit rate which moved in the expected direction was that for guns from vehicle stops which rose 1.8 percentage points at the pandemic onset as overall frisks dropped. The pandemic findings are difficult to interpret for several reasons. Among them, predictions from policing models using a suspicion threshold are ambiguous for this event. This is due to the fact that frisk rates conditional on stop actually went up during this period, so hit rate predictions are ambiguous. We discuss further impediments to learning from this event below.

Another way to examine changes in policing is by looking at frisks and searches as a share of all stops. A large drop in stops should have the same implication as a large decline in frisk - it should increase the stop hit rate, namely the share of stops in which contraband is discovered. We examine this in Table 3, which reports results from estimating equation 1 on data from 2015 to 2020 for all stops using both the protests and pandemic events. The results here are similar to that in Table 2, although all point estimates are now positive and a larger proportion are statistically significant.

While the difference-in-difference approach in Tables 2 and 3 is preferred, because it accounts for seasonality in the most fine-grained way, one may be concerned that prior year data is driving the results. As such we run a single difference (before-after) comparison of hit rates around the protest, only using 2020 data. The results (Appendix Table A2) are again consistent although the magnitudes of the impact are generally slightly larger. The increases in contraband recovery from pedestrian stops in Chicago and the increases in both all contraband and gun hit rates from vehicles stops in Philadelphia are now also statistically significant at the 5% level.

Figure 3 shows results from the following event study specification:

$$contraband_i = \alpha + \sum_{t=-6}^{11}(\beta_{1t} * Week_{it} * Treat_i) + \sum_{t=-6}^{11}(\beta_{2t} * Week_{it}) + \sum_{k=2015}^{2020}(\gamma_k * Year_{ik}) + \sum_{j=1}^{66}(\delta_j * PSA_{ij}) + \theta_i + \epsilon_i \quad (2)$$

Where variables are defined the same as in equation 1. $Week_{it}$ is a week dummy for the number of weeks which the stop occurred after 29 May for the given year for Chicago and 3 June for Philadelphia. For the event study we focus on the protest event because as we have seen in Figure 1, the decline in frisks was lower and briefer during the pandemic onset. The time span in Figure 3 is larger than that in our main difference-in-difference specification in Table 1. What seems clear is that the hit rate from pedestrian frisks in Chicago and vehicle searches in Philadelphia increased substantially after the protests and seemed to largely stay at the higher level for most of the 3 months of post-protest data in the figure. The picture is less clear for pedestrian frisks in Philadelphia, which are noisier.

While most of these initial results are consistent with expectations from standard models of police stops, there may be concern that other changes could impact the findings. 2020 was certainly not a typical year and the time periods studied were specifically chosen because there was a marked shift in policing. It is likely that other relevant changes occurred at the same time that could explain the findings. There are several other reasons why hit rate may have increased as stops and frisks declined. These include a change in the number or composition of potential suspects, change in police deployment, or a change in police effort. We explore each of these possibilities in turn.

A. Testing for Changes in Suspect Population

The number of potential suspects on the street or in vehicles could increase or decrease hit rate depending on one's model of officers. A greater number increases potential suspects, so if observable characteristics are informative of hit rate officers should have a larger pool of higher likelihood suspects and thus a greater hit rate. However, if an increased suspect pool goes along with changing composition, such that most of the additional traffic is unlikely to have contraband, and officers only have weak predictive ability of contraband carriers based on observables, then an increase in potential suspects could decrease the hit rate.

Figure 4 displays two indices from Google Community Mobility Reports, retail and recreation in the blue solid line and transit in the red dashed line. These provide a measure of foot traffic and follow a similar pattern in both cities, with a sharp, deep drop around the pandemic onset and then a slow rise beginning in April, 2020. This suggests that any attempt to draw inference about the decline in police stops from the pandemic period will be impeded by this substantial simultaneous change. Given the discussion above this would mean that the estimates of the hit rate change could be either upward or downward biased, depending on our model of policing and assumptions about the characteristics of additional people on the street. This does not apply to the period around the protests when there was a smooth but relatively modest increase in mobility measures.

Evidence on changes in vehicular travel comes from age breakdowns of motor vehicle crash patients as recorded by the CPD in Chicago and in the PTOS trauma database in Philadelphia. These are plotted in Figure 5. The top panel for each city shows changes around the pandemic onset and the bottom around the protests. In each figure the blue dashed bars measure the density prior to the event and the red solid bars afterward. There is substantial change around the pandemic with a substantial tightening in the age distribution in Philadelphia. The change around protests is smaller for both cities, although in Philadelphia there is a decrease in the share of children and an increase in individuals in their 20's and early 30's after the protests. A χ^2 test was run to test for changes in the age distribution around the protests (Appendix Table A3) and in both cities the test does not reject the null of no change.

Taken together it seems likely that the potential suspect pool changed markedly around pandemic onset. For this reason we focus on changes around the protests for the rest of the paper. Even during this period, there were certainly localized changes in potential detainees as protests or looting occurred. But overall mobility patterns changed smoothly and one measure of vehicular travel didn't change appreciably. We further explore local variation around the time of the protests below.

B. Testing for Changes in Police Deployment

An important implication of the policing models is that as an officer decreases the number of stops she makes, the hit rate on the marginal stop should increase. But the decline in police stops could result from a decline in officers making stops, a decline in stops per officer or both. Depending on spatial allocation of officers, these could have different theoretical implications for hit rates. For example, if officers patrol a set area and overlap with others who are redeployed elsewhere during the protests, they may be expected to encounter more suspects and hence have a higher hit rate. If patrol territory doesn't overlap others and suspects stay within a single patrol area one might expect no impact on hit rate on average.

We do not have information on officer patrol area, but we can observe consistent identifiers of officers making stops. Thus we are able to determine whether the decline is on the intensive margin - stops per officer, or the extensive margin - number of officers making stops. Figure 6 reveals that during the protests, there was a decline on both margins, as the solid black and dotted red lines plummet in both Chicago and Philadelphia. In both cities the shift in the extensive margin is greater. In Philadelphia the number of officer pairs making stops fell by about 75% while the decline in stops per officer pair was about 35%. The change in Philadelphia during the pandemic is particularly interesting, with no decline in the intensive margin, but a large drop on the extensive margin.

While there was a large decline in stops during both periods, the difference in stops per officer pair could explain the differing results in Table 2, which showed that hit rates mostly declined during the pandemic but rose during protests. Shifts in the extensive margin could have an ambiguous impact on hit rates, depending on deployment of officers and distribution of suspects. But declines on the intensive margin should clearly increase hit rates - what we see during the protests.

One way to isolate the impact of the change in intensive margin is by focusing on a subset of officers who made stops in the 6 weeks prior to the protests. Table 4 reports results from these regressions, which are similar to the main results. Hit rates generally increase by a similar magnitude as in the main specifications reported in Table 2. This is an essential result to be consistent with the policing models: one would expect as individual officers become more selective in stops and frisks, their hit rates will increase.

One concern about the sharp decline in stops per officer pair in Figure 6 is that this may indicate a reduction in police effort. If that were the case, one would expect a decline in hit rate, or at best a modest increase. It is not possible to completely rule out the possibility that effort fell during the protest and that the increase in hit rate would otherwise be even greater. One additional piece of information that casts doubt on the reduced effort possibility is provided in Table 5. This table reports results from estimation equation 1, but where the dependent variable is a dummy for whether the stop or frisk was legally unfounded. In both cases, the point estimates show a decline in illegal stops and frisks, which is one measure of police officer effort. Thus, it seems unlikely that large changes in police effort have had a big impact on the results.

C. Change in Crime

While we have examined potential changes in the composition of pedestrians and drivers in section A above, it is still possible that there could be a change in criminal propensity not detectable in numbers or demographics. If the share of potential criminals increased in the suspect population, one would expect hit rates to rise, even absent any change in police numbers.

To investigate this, we estimate equation 1 using contemporaneous crime reports as the dependent variable (Table 6). While the change in crime varies substantially by crime type, most policing tends to be based on violent crime. Column 2 of Panel B shows no significant change in violent crime in Philadelphia over the period of the protests. In fact, column 1 shows a substantial decrease in crime overall, which makes it highly unlikely that changes in criminal propensity among the suspect population is responsible for the rise in hit rates in Philadelphia. If anything, criminality may have fallen slightly in this period, without which the hit rate increase may have risen.

In contrast, in Chicago there was a large and statistically significant increase in most types of violent crime in the period after the protests. Overall violent crime increased by 20%

and shootings increased by around 45%. Given that an increase in shootings should serve as a proxy for the gun carrying rate, this increase would be expected to increase hit rates even in the absence of a change in police behavior. For example, under a model of random search an increase in the gun carry rate by 10% would increase the hit rate by 10%. To investigate whether an increase in criminality drove the increase in hit rates, we estimate the change in hit rates for a subset of the city where shootings did not increase significantly. The 8 police sectors we look at are shown in black in Figure 7 and consist of an almost contiguous subset of the city representing roughly half of the police stops. The change in crime following the protests in this subset of Chicago can be seen in Panel A of Table 7. Across these sectors the changes to overall violent crime and shootings specifically were much smaller in magnitude (only 10% and 5.5% respectively) and statistically insignificant. Despite this the increase in the hit rate for guns from pedestrian frisks is of a larger magnitude with a 4 percentage point increase compared to 2.5 percentage points for the whole of Chicago. However, given the lower number of observations leading to an increase in the standard error it is no longer statistically significant. The increase in general contraband hit rates from both vehicle searches and pedestrian frisks is also larger in magnitude than for Chicago overall.

D. Race Effects

To this point, this analysis has focused on overall changes in hit rates. But hit rate tests are most frequently employed to assess racial disparities in policing, and it is worthwhile to examine what occurred during the time periods under study. In the discussion of Table 1, we noted the fairly substantial decline in the Black share of stops in Philadelphia during the protests, both pedestrian and vehicle. It is certainly possible, if not likely, that the subject of the protests - racially disparate policing - was a driver of this change. This seems even more likely when comparing with the pandemic onset which saw a substantial rise in the minority share of stops.

$$contraband_i = \alpha + \beta_1 after_i * treat_i * Black_i + \beta_2 after_i * treat_i + \beta_3 after_i * Black_i + \beta_4 treat_i * Black_i + \beta_5 Black_i + \beta_6 after_i + \sum_{k=2015}^{2020} (\gamma_k * year_{ik}) + \sum_{j=1}^{66} \left(\delta_j * PSA_{ij}\right) + \theta_i + \epsilon_i \tag{3}$$

If reaction to the protests can be understood as increasing the threshold for stopping or frisking a Black pedestrian or motorist, then we should expect an increase in hit rate for Blacks relative to whites. Table 8 reports results from the triple difference specification in

The results here are mixed. Even as the share of Blacks stopped declines, the relative hit rate also drops appreciably for vehicle stops. But for pedestrian stops the change is quite small for gun hit rate and positive and large for all contraband, as would be expected. This may suggest a couple of possibilities. First, the underlying contraband carrying rate among Blacks in vehicles may have declined relative to whites during this period of uncertainty around the protests. Second, due to the protests, police officers may have been more likely to simply issue a warning and hence not record as many cases of contraband for Blacks relative to whites, although why this would apply simply to vehicles is unclear.

V. Additional Results and Discussion

A. Robustness Checks

Choosing the correct window to evaluate the impact of events like those under study here always involves a tradeoff. A short window decreases the likelihood that other changes contaminate the natural experiment but comes along with less data. A longer window increases the number of observations but weakens the focus. In Table A4, we present results with varying time windows to assess the sensitivity of the results to precise timing.

Two other timing issues are examined, besides just the size of the window. In columns 1-4 we consider a much longer "after" period - 12 weeks and find generally consistent results. This is important to address the real possibility that it takes some time for individuals to adjust to a new equilibrium of lower stop and frisk rates - both officers and suspects. The findings in these columns suggest that much of the adjustment occurs within the first 6 weeks after the protests, the "after" period used in most regressions.

In Philadelphia, the decline in stops and frisks did not occur immediately when the protests began, but about 4 days later. During the first week of protests the disruptive effects of the protests on individual behavior are also likely greatest. Thus, for both cities in columns 5-8 we exclude the first week of protests from the analysis. We find almost identical results for Chicago while results in Philadelphia are also similar but of larger magnitude. Finally, in columns 9-12, we exclude a period of 2 weeks after the first protest during which most of the additional protests occurred and find the results again to be consistent.

B. Discussion

While most of the main specifications show an increased hit rate while stops and frisks dropped tremendously around the protests, the hit rate declines for pedestrian stops in in Philadelphia the all contraband specification, which is driven by drugs. At the same time it rises substantially for vehicle stops and all contraband. These results may actually be consistent with a greater emphasis by the police on strictly adhering to the law in stops and frisks in response to the protests.

As mentioned above, a legal pedestrian stop requires that a police officer has reasonable suspicion that the person to be stopped is or is about to be engaged in criminal activity. To then frisk the individual, the officer must have reasonable suspicion that they are armed. Importantly, a suspicion that the individual is carrying drugs is insufficient grounds for a frisk. Given the significant public scrutiny of police behavior at the time of the Protests, it would be expected that police were exercising more care to ensure the legality of their actions and as seen in Table 5 the proportion of illegal frisks decreased. Hence, we would expect that police reduced the number of frisks conducted solely for the purpose of discovering drugs (which would constitute an illegal frisk). Given drugs make up the vast majority of contraband recovered from police stops this would lead to a reduction in the hit rate for pedestrian stops with respect to overall contraband recovered, consistent with what we see in column 6 of Table 2.

For vehicle stops, a police officer can stop any vehicle where the driver or occupant is observed violating the law (or where the officer reasonably believes they were violating the law) and the vast majority are for traffic violations. The requirement of reasonable suspicion that the that the individual is armed is the same for frisks of the driver/occupants of the vehicle as for pedestrian stops. However, the probable cause standard for a vehicle search can be easily satisfied by suspicion of any criminality, including a drug violation based on an odor emanating from the vehicle. Hence a greater focus on the legality of police activity would not necessarily have the same impact on the recovery of drugs as in the case of pedestrian stops, and thus the drop in overall search rate still dominates, resulting in the higher hit rate.

Before concluding, it is worth considering the magnitude of the hit rate response to the large decline in stops. A simple calculation shows that in Philadelphia for gun hit rates a 10% rise in stops corresponds to a 0.18 percentage point decline in pedestrian hit rate and 0.22 percentage point decline in vehicle hit rate. The change in hit rate with respect to stops or frisks is directly related to the second derivative of the guns-stops relationship. To make optimal policy decisions with respect to police stops, one needs to know not only the hit rate, which is easy to compute, but information about the second derivative, which hasn't been previously explored.

VI. Conclusion

If not for the biggest pandemic in a century, the role of race in policing would have been the dominant news story in the year 2020. It is a topic that recurs with increasing force and urgency and we attempt to add to our understanding of it. We take advantage of the drastic reductions in pedestrian and vehicle police stops, frisks, and searches following the pandemic onset and nationwide protests. We provide empirical corroboration of the salient predictions of optimizing models of police behavior: the contraband hit rate should rise when the number of stops per officer falls, *ceteris paribus*.

Indeed, we find that hit rates from pedestrian and vehicle stops generally rose as stops and frisks fell dramatically. Importantly, with detailed complementary data, we are able to rule out a number of alternative explanations, including changes in street population, crime, police allocation, and policing intensity. In addition, we find mixed evidence about the changes in racial disparities. The results are robust to a number of different specifications. While an increase in the hit rate is implied by both the KPT and Anwar and Fang models, given the large increase seen in such a short time frame we believe it is unlikely to be driven purely by a change in driver behavior in response to the lower probability of detection. Hence our results appear to favor the model of Anwar and Fang.

Our findings have important implications for potential reforms to improve policing in the United States. First, policing is a very noisy process, where the vast majority of the searches/frisks do not result in contraband findings. This suggests that effective policing can benefit greatly from more community and neighborhood engagement, so that police can make decisions about search/frisks with more accurate information. The police can also benefit from more training about best practices to identify guilty subjects. This could lead to fewer tense confrontations between police and the citizens.

Second, despite the admittedly noisy policing process, the findings in our paper also suggest that police behavior is broadly consistent with models where they aim to at least partly maximize the contraband finding rates, using the noisy and imperfect signals they have at the time of making their decisions.

References

Abrams, D. S., Bertrand, M., & Mullainathan, S. (2012). Do Judges Vary in Their Treatment of Race? *Journal of Legal Studies*, *41*(2), 347–383.

Antonovics, K., & Knight, B. G. (2009). A New Look at Racial Profiling: Evidence from the Boston Police Department. *The Review of Economics and Statistics*, *91*(February), 163–177. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/rest.91.1.163

Anwar, S., & Fang, H. (2006). An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence. *The American Economic Review*, *96*(1), 127–151. Retrieved from http://www.nber.org/papers/w11264

Ayres, I. (2002). Outcome Tests of Racial Disparities in Police Practices. *Justice Research and Policy*, *4*(SI), 131–142. https://doi.org/10.3818/JRP.4.1.2002.131

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *The Journal of Political Economy*, *76*(2), 169–217. Retrieved from http://www.journals.uchicago.edu/doi/abs/10.1086/259394

Bjerk, D. (2007). Racial Profiling, Statistical Discrimination, and the Effect of a Colorblind Policy on the Crime Rate. *Journal of Public Economic Theory*, *9*(3), 521–545. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9779.2007.00318.x/full

Devi, T., & Fryer Jr., R. G. (2020). Policing the Police: The Impact of “Pattern-or-Practice” Investigations on Crime. In *National Bureau of Economic Research (NBER)*. Retrieved from http://www.nber.org/papers/w27324

Dharmapala, D., & Ross, S. L. (2004). Racial Bias in Motor Vehicle Searches: Additional Theory and Evidence. *Contributions to Economic Analysis & Policy*, *3*(1), Art. 12. Retrieved from http://www.degruyter.com/view/j/bejeap.2003.3.issue-1/bejeap.2004.3.1.1310/bejeap.2004.3.1.1310.xml

Durlauf, S. N. (2006). Assessing Racial Profiling. *The Economic Journal*, *116*(515), F402–F426. https://doi.org/10.1111/j.1468-0297.2006.01129.x

Feigenberg B, Miller C. (2020) Racial disparities in motor vehicle searches cannot be justified by efficiency. *National Bureau of Economic Research*; 2020 .

Gelbach, J. B. (2021). Testing Economic Models of Discrimination in Criminal Justice. *Available at SSRN 3784953*

Heaton, P. (2010). Understanding the Effects of Antiprofiling Policies. *Journal of Law & Economics*, *53*(1), 29–64. Retrieved from https://proxy.library.upenn.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cja&AN=52445028&site=ehost-live

Knowles, J., Persico, N., & Todd, P. (2001). Racial Bias in Motor Vehicle Searches: Theory and Evidence. *Journal of Political Economy*, *109*(1), 203–229.

Knox, D., & Mummolo, J. (2020). Toward a General Causal Framework for the Study of Racial Bias in Policing. *Journal of Political Institutions and Political Economy*, *1*, 1–38. https://doi.org/10.1561/113.00000018

Melamed, S., & Newall, M. (2020). With Courts Closed by Pandemic, Philly Police Stop Low-Level Arrests to Manage Jail Crowding. *The Philadelphia Inquirer*, pp. 1–7.

Persico, N. (2002). Racial Profiling, Fairness, and Effectiveness of Policing. *The American Economic Review*, *92*(5), 1472–1497. Retrieved from http://www.ingentaconnect.com/content/aea/aer/2002/00000092/00000005/art00010

Ridgeway, G. (2006). Assessing the Effect of Race Bias in Post-traffic Stop Outcomes Using Propensity Scores. *Journal of Quantitative Criminology*, *22*(1), 1–29. https://doi.org/10.1007/s10940-005-9000-9

Sanga, S. (2009). Reconsidering Racial Bias in Motor Vehicle Searches: Theory and

Evidence. *Journal of Political Economy*, *117*(6), 1155–1159.
https://doi.org/10.1086/318603

Figures

Figure 1a

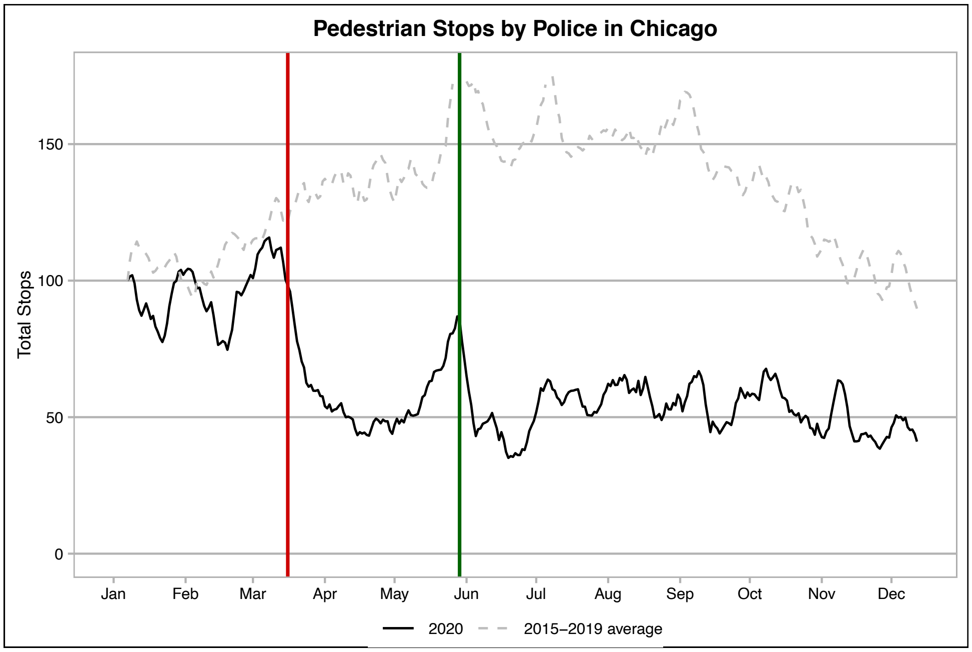
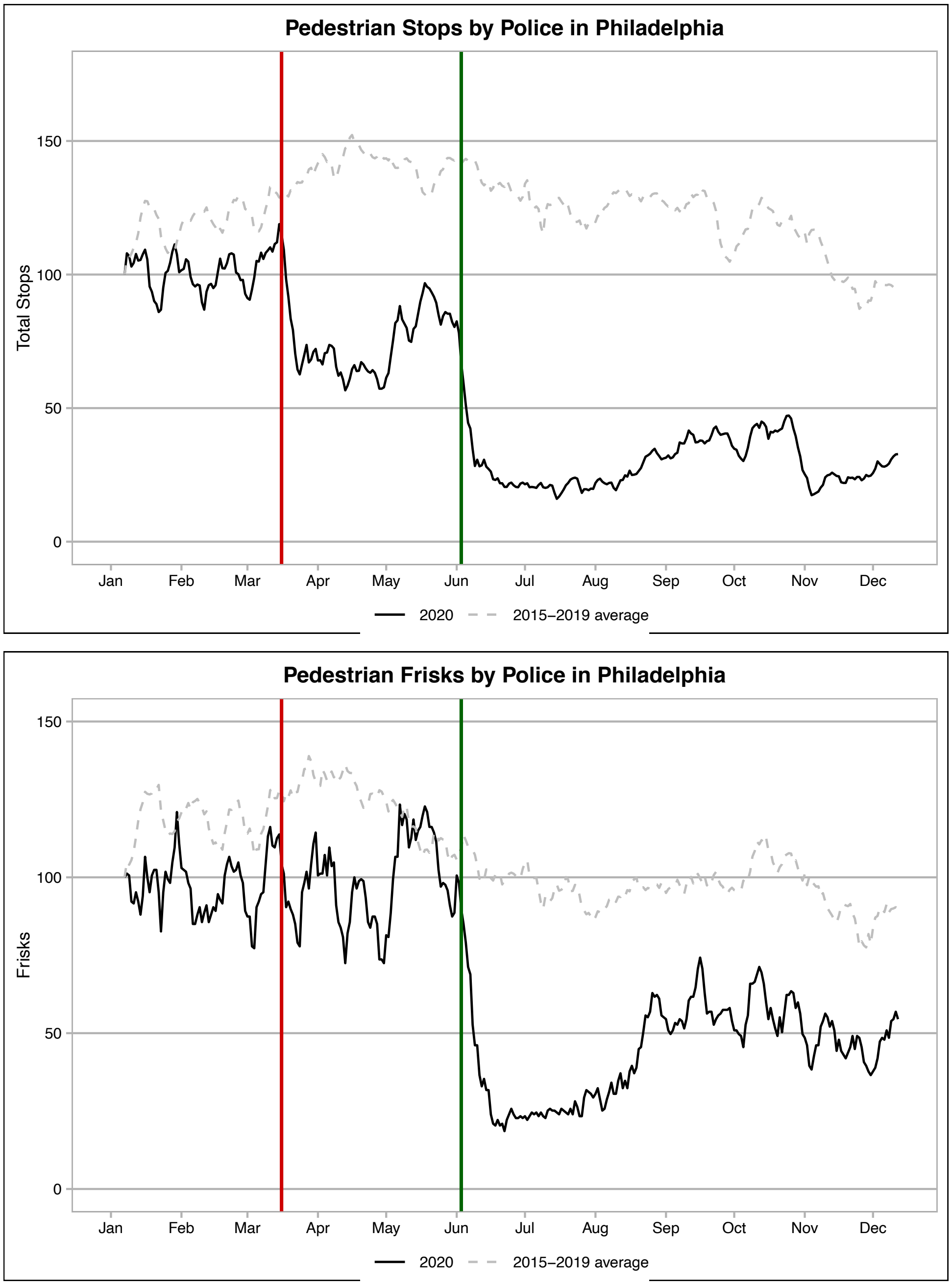
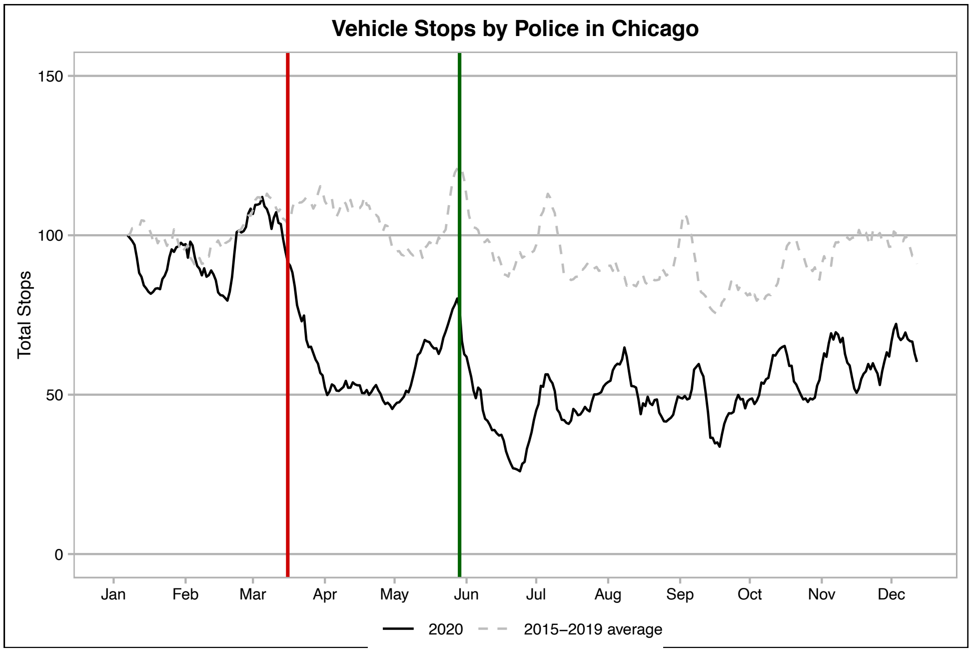


Figure 1b



Panel A shows the number of Police stops of pedestrians in Chicago while Panel B shows the number of frisks of pedestrians. Panels C and D show the equivalent for Philadelphia. The data is shown for 2020 (black) and the average for the prior years (grey). Both series are indexed to the average for the week from 1-7 January of the relevant year(s). The red vertical line indicates the onset of the pandemic while the green indicates the fall in stops in response to the George Floyd protests. Data Source: Philadelphia Police Department & Chicago Police Department

Figure 2a


Vehicle Stops by Police in Chicago
150
100
50
0
Total Stops
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
2020
2015–2019 average

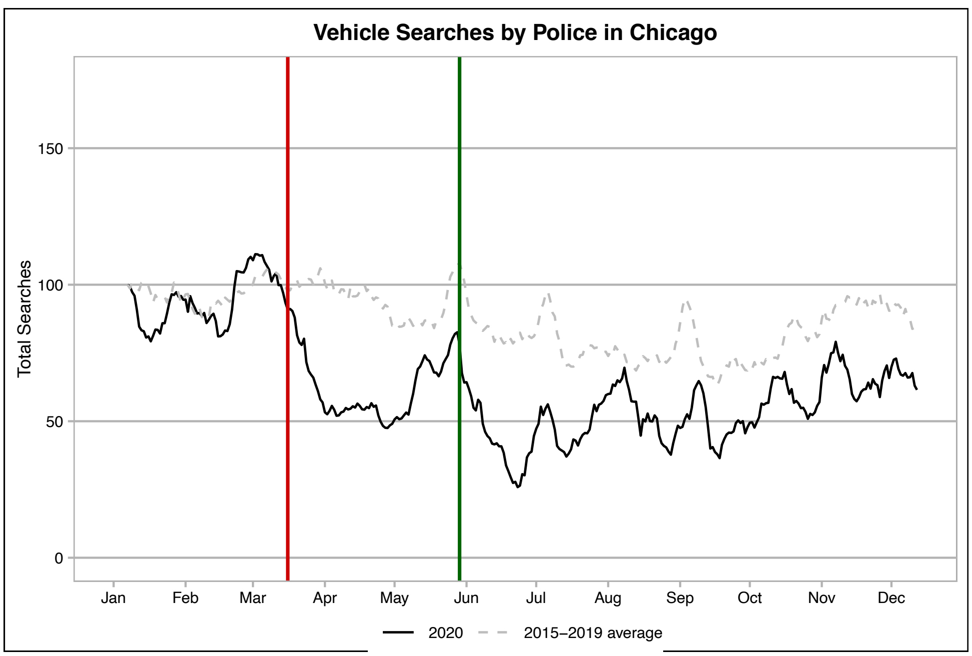
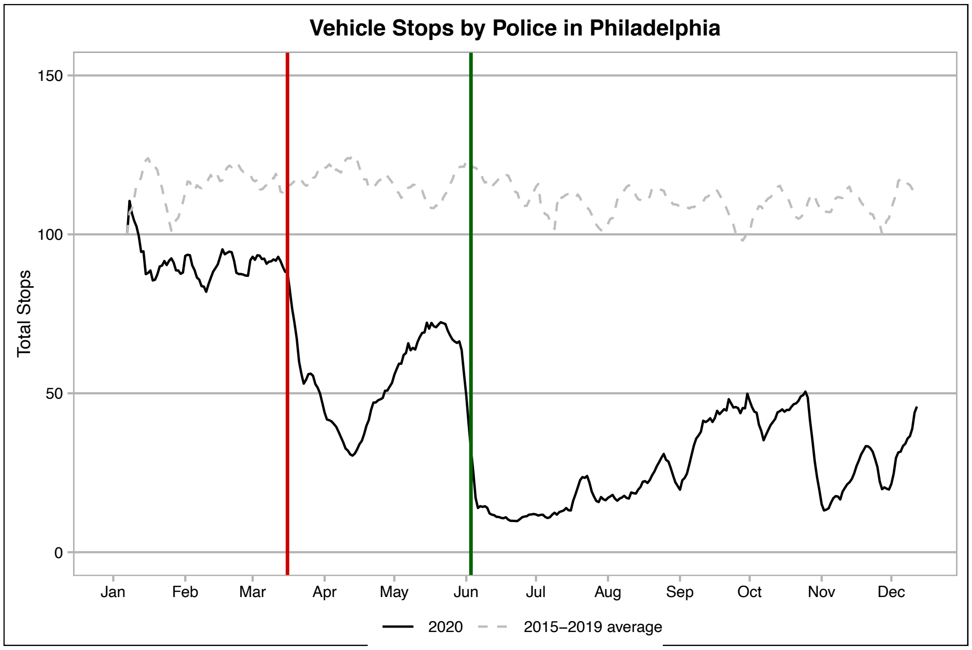
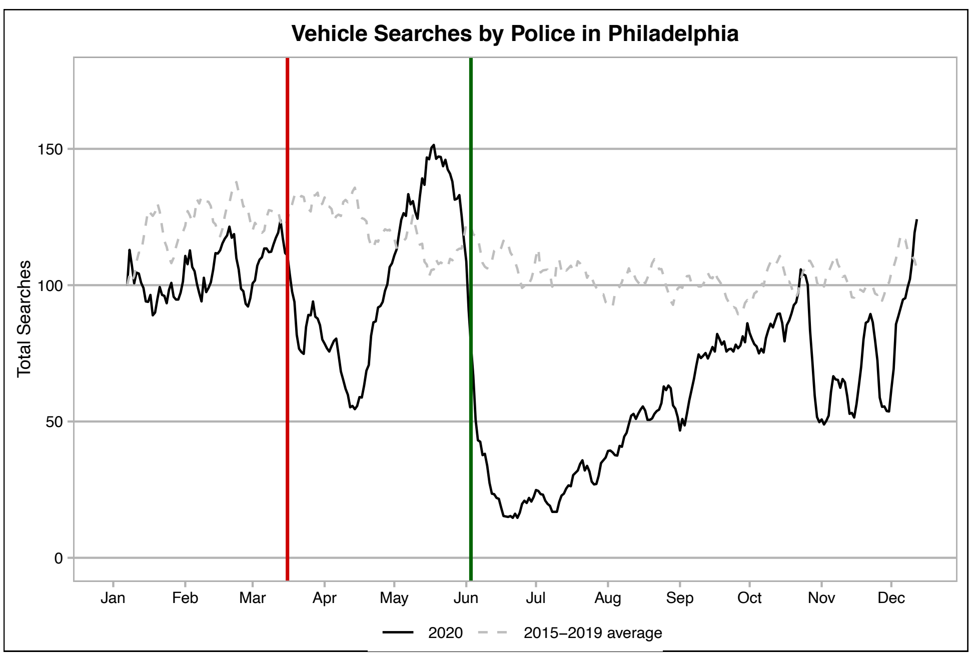

Vehicle Searches by Police in Chicago
150
100
50
0
Total Searches
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
2020
2015–2019 average

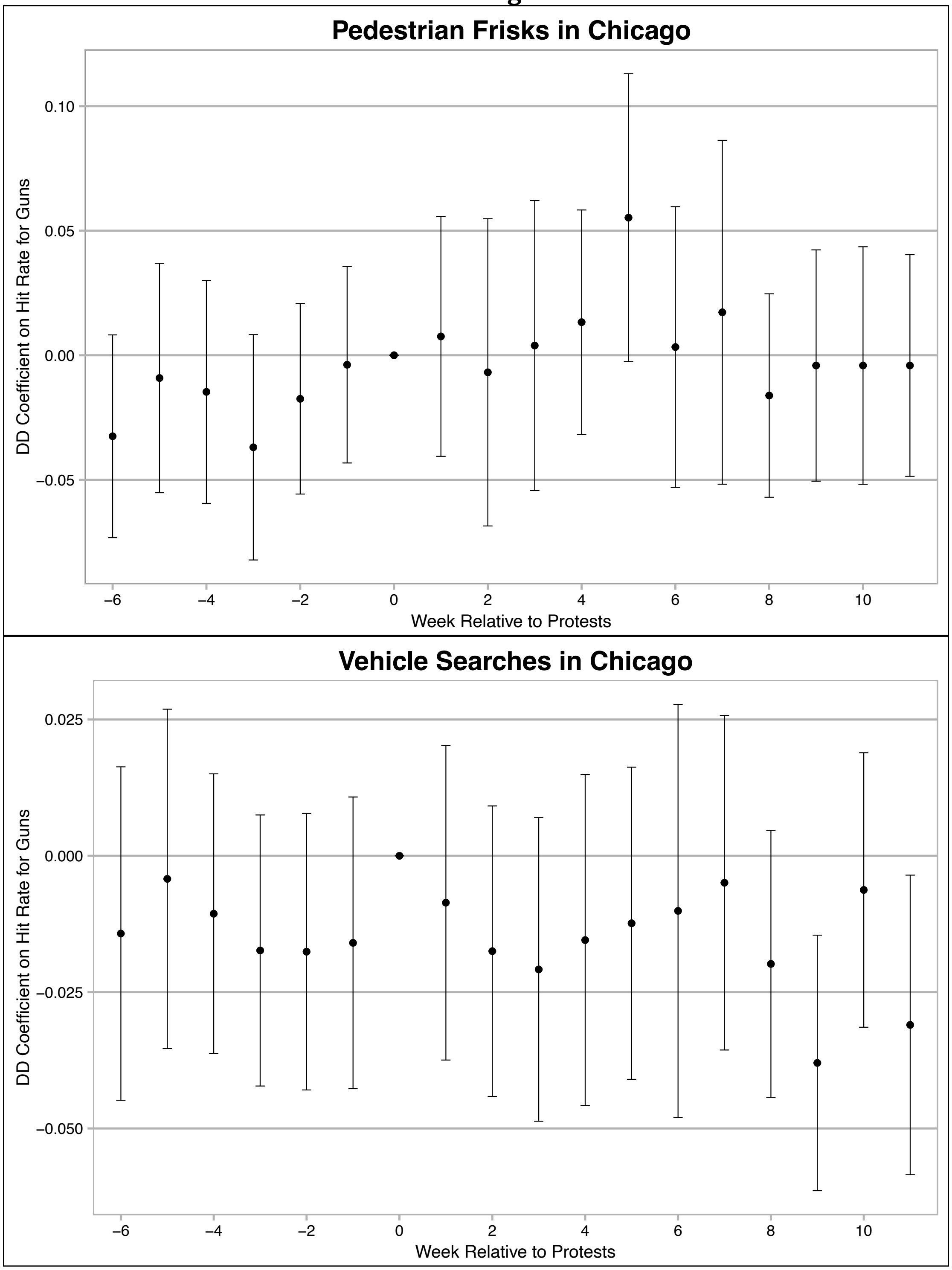
Figure 2b

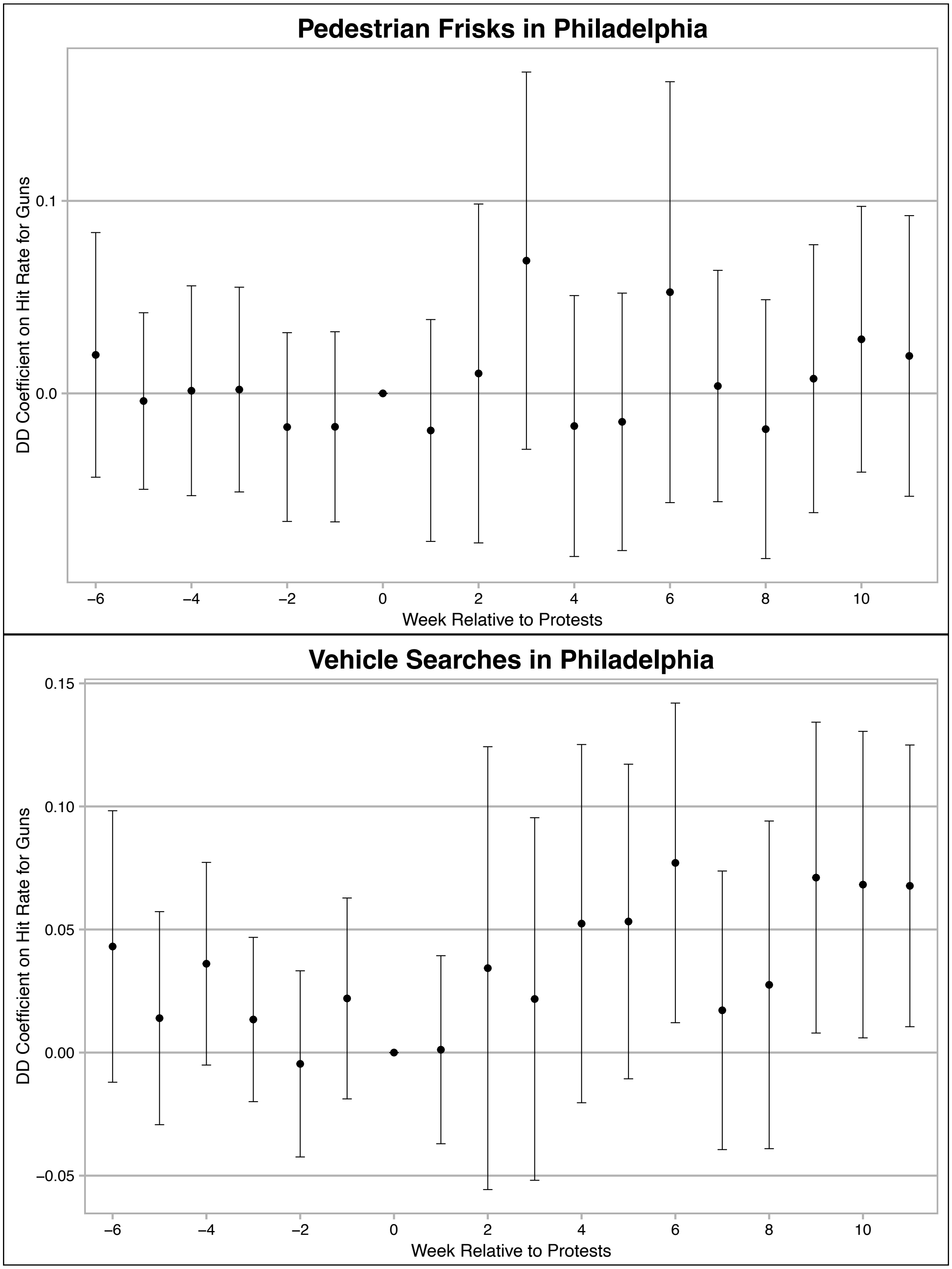




Panel A shows the number of vehicle stops in Chicago while Panel B shows the number of vehicle searches. Panels C and D show the equivalent for Philadelphia. The data is shown for 2020 (black) and the average for the prior years (grey). Both series are indexed to the average for the week from 1-7 January of the relevant year(s). The red vertical line indicates the onset of the pandemic while the green indicates the fall in stops in response to the George Floyd protests. Data Source: Philadelphia Police Department & Chicago Police Department

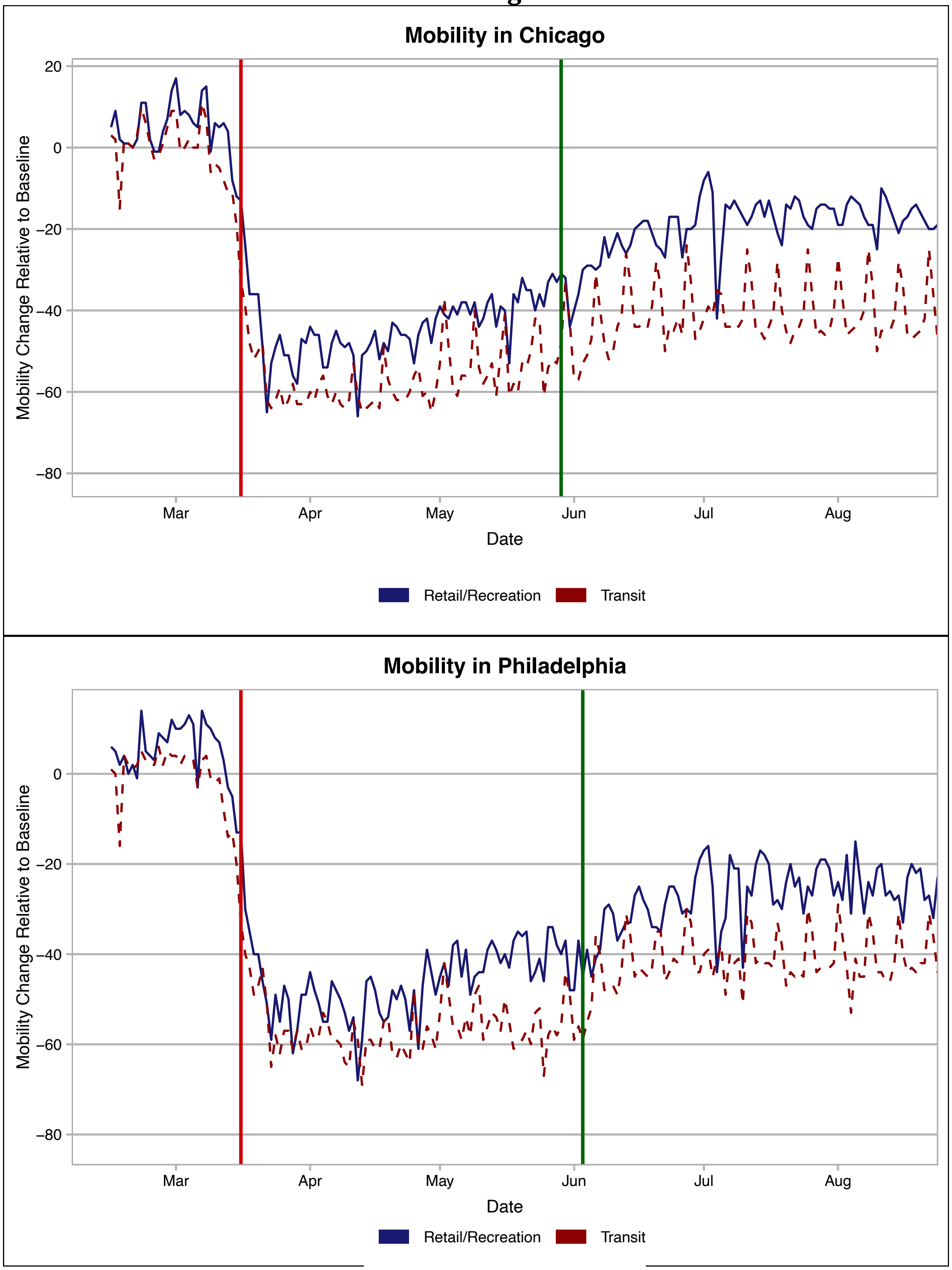
Figure 3


Pedestrian Frisks in Chicago
DD Coefficient on Hit Rate for Guns
0.10
0.05
0.00
−0.05
−6
−4
−2
0
2
4
6
8
10
Week Relative to Protests
Vehicle Searches in Chicago
DD Coefficient on Hit Rate for Guns
0.025
0.000
−0.025
−0.050
−6
−4
−2
0
2
4
6
8
10
Week Relative to Protests



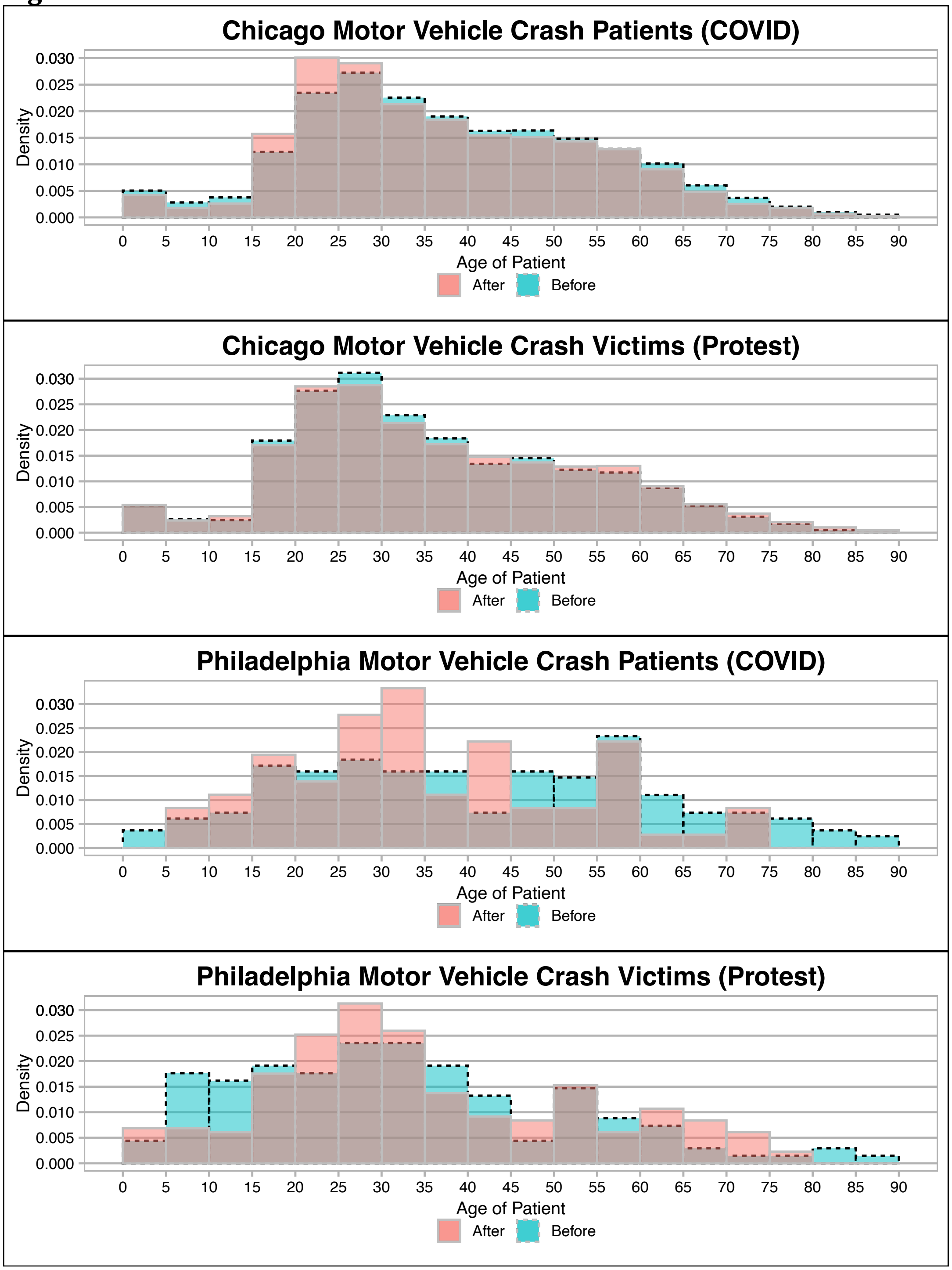
Event study: Hit Rate for finding guns conditional on frisk/search in 2020 relative to 2015-2019. The figures show the plots of the regression coefficients from OLS of guns on dummies for week of 2020. Specification is estimated on data from the days 6 weeks before to 12 weeks after the George Floyd protests, along with the same calendar dates for 2015-2019. Year fixed effects are included as well as controls for age, gender, race and police region in which the stop occurred. The vertical lines for each coefficient show the 95% confidence intervals from robust standard errors clustered at the region level

Figure 4



Shows the change in mobility in Chicago and Philadelphia over 2020 relative to a baseline established Jan 3 – Feb 6 2020. Mobility is reported for the Retail and Recreation (blue solid line) and Transit (red dashed line) categories reported by the Google Community Mobility Reports. The red vertical line indicates the Pandemic onset while the green indicates the fall in stops in response to the George Floyd protests. Data Source: Google Community Mobility Reports

Figure 5

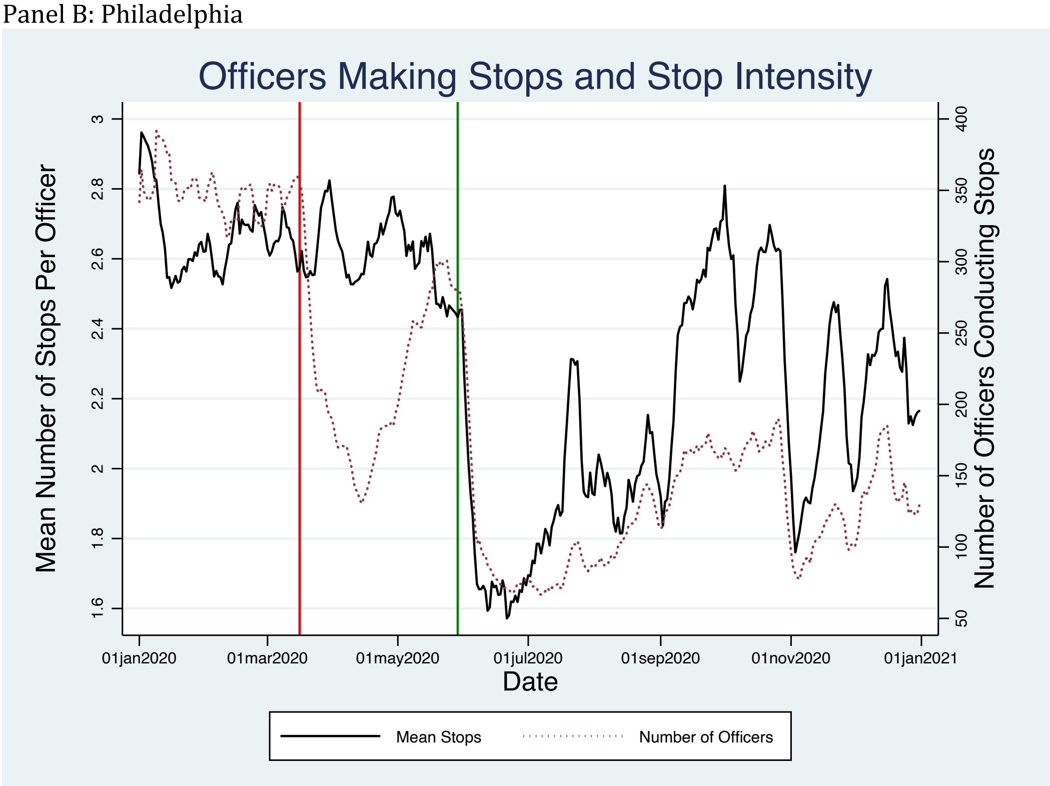


Density histogram showing the age distribution of motor vehicle crash victims before and after each event in Chicago and Philadelphia. The first plot for each city is relative to the onset of the COVID-19 pandemic while the second plot is relative to the George Floyd Protests. Data from 6 weeks before to 4 weeks after the pandemic are used and 3 weeks before to 6 weeks after the George Floyd Protests. Period relative to the event is indicated by the color and outline of the bars, red with a solid outline being the days after and blue with a dashed outline the days before. Data Source: Pennsylvania Trauma Outcomes Study (PTOS) & Open Data Chicago

Figure 6

Panel A: Chicago

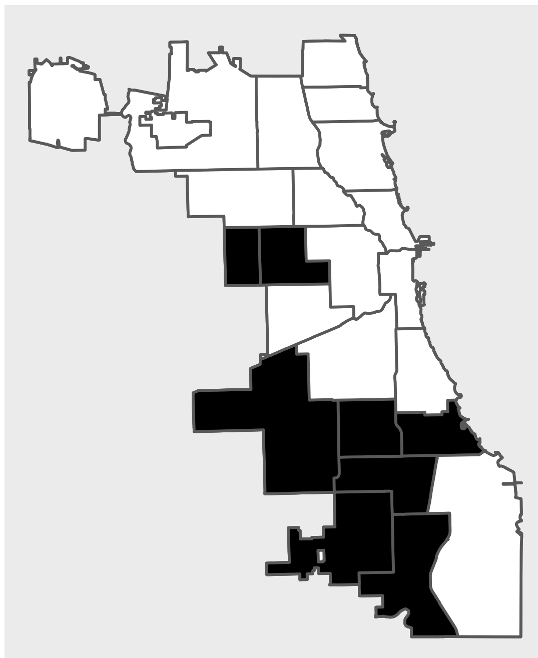
Panel B: Philadelphia



The left axis indicates the scale for the average number of stops conducted by officers who conduct at least one stop on a given day, shown in black. Panel A shows the data for Chicago while Panel B shows the data for Philadelphia. The red vertical line indicates the Pandemic onset while the green indicates the fall in stops in response to the George Floyd protests. Both series are smoothed with a 7 day moving average. Data Source: Philadelphia Police Department & Chicago Police Department

Figure 7

Selected Sectors of Chicago



Map of Chicago. Solid black lines mark the boundaries of each police sector. The 8 sectors shaded in black are among the districts with the lowest increase in shootings in the period following the 2020 protests. These districts are the ones selected as a subset of Chicago for separate analysis.

Tables

Table 1: Summary Statistics

	Pedestrian Stops				Vehicle Stops			
	Chicago		Philadelphia		Chicago		Philadelphia	
	Before	After	Before	After	Before	After	Before	After
Panel A: Pandemic								
Stops per Day	228	89	118	78	208	78	784	393
Frisks per Day	57	29	23	22	113	47	89	62
% Male	88%	87%	86%	89%	82%	83%	72%	75%
% Black	65%	64%	70%	78%	64%	64%	74%	80%
Age	35.7	34.5	33.9	33.3	28.3	26.7	34.9	33.0
Contraband|Frisk/Search=1	13.3%	14.9%	12.1%	10.2%	23.9%	23.7%	19.5%	18.5%
Contraband|Race=Black,Frisk/Search=1	13.8%	13.9%	12.2%	9.3%	23.6%	22.0%	18.2%	18.3%
Contraband|Race=Other,Frisk/Search=1	12.1%	17.3%	12.0%	13.3%	24.5%	27.9%	24.1%	19.5%
Gun|Frisk/Search=1	2.4%	4.3%	3.1%	2.8%	1.3%	2.0%	3.3%	5.0%
Gun|Race=Black,Frisk/Search=1	2.6%	4.0%	3.4%	2.5%	1.6%	2.1%	3.4%	5.1%
Gun|Race=Other,Frisk/Search=1	1.9%	5.1%	1.9%	3.7%	0.6%	1.8%	2.7%	4.9%
Panel B: Protests								
Stops per Day	126	71	99	26	138	75	522	105
Frisks per Day	42	26	25	7	84	44	105	19
% Male	90%	87%	87%	85%	86%	85%	78%	77%
% Black	63%	62%	76%	71%	63%	70%	79%	75%
Age	34.4	33.6	32.7	35.5	26.5	27.0	32.1	33.3
Contraband|Frisk/Search=1	14.2%	19.2%	14.0%	12.8%	25.4%	29.2%	21.1%	24.6%
Contraband|Race=Black,Frisk/Search=1	16.0%	20.4%	13.4%	13.4%	27.4%	30.6%	20.9%	24.0%
Contraband|Race=Other,Frisk/Search=1	10.7%	16.5%	16.0%	10.9%	21.2%	25.4%	22.1%	26.7%
Gun |Frisk/Search=1	5.8%	9.1%	2.8%	4.4%	2.5%	3.0%	3.7%	5.6%
Gun |Race=Black,Frisk/Search=1	7.8%	10.3%	3.2%	4.8%	3.2%	3.6%	3.9%	5.3%
Gun|Race=Other,Frisk/Search=1	2.0%	6.0%	1.7%	3.1%	1.1%	1.3%	3.1%	6.4%

Police investigation summary statistics. For the stay-at-home orders, "Before" is the 6 weeks before 16 March 2020 and "After" is the 4 weeks directly after. For the George Floyd Protests, "Before" is the 3 weeks before the pandemic began in the relevant city and "After" is the following 6 weeks. Race and gender are reported as determined by police at time of stop. Contraband is whether any contraband is flagged by the police (includes drugs and weapons) and reported as a proportion of total frisks/searches conducted. "Gun" reflects whether the contraband found was a firearm and was manually coded based upon the police description of results from each investigation. Data Source: Philadelphia Police Department & Chicago Police Department

Table 2A: Impact of Protests on Frisk/Search Hit Rate

	Chicago				Philadelphia			
	All Contraband		Guns only		All Contraband		Guns only	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After*Treat	0.034*	0.031	0.004	0.025*	0.043	-0.015	0.017	0.013
	(0.017)	(0.017)	(0.006)	(0.011)	(0.023)	(0.028)	(0.013)	(0.016)
After	0.006	0.014*	0.001	0.008**	-0.006	0.001	0.003	9.00E-04
	(0.01)	(0.006)	(0.002)	(0.003)	(0.007)	(0.005)	(0.004)	(0.002)
Treat	0.114**	0.065**	0.014*	0.049**	0.073**	0.034*	0.012	0.018*
	(0.021)	(0.015)	(0.006)	(0.009)	(0.016)	(0.017)	(0.007)	(0.008)
Black	-0.020*	-0.005	0.007**	-0.0002	-0.035**	-0.008	0.007	0.003
	(0.01)	(0.009)	(0.002)	(0.004)	(0.007)	(0.008)	(0.004)	(0.005)
Observations	20,024	16,794	20,024	16,794	25,507	15,041	25,507	15,041
Adjusted R2	0.020	0.014	0.006	0.017	0.024	0.006	0.006	0.006
Mean Y	0.195	0.112	0.016	0.029	0.148	0.108	0.031	0.018

Note: $^{*}p<0.05$ $^{**}p<0.01$

Table 2B: Impact of Pandemic Response on Frisk/Search Hit Rate

	Chicago				Philadelphia			
	All Contraband		Guns only		All Contraband		Guns only	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After*Treat	0.011	0.030	0.006	0.021*	-0.002	-0.015	0.018*	-0.004
	(0.016)	(0.017)	(0.005)	(0.009)	(0.019)	(0.021)	(0.008)	(0.009)
After	-0.003	-0.009	0.001	-0.002	-0.003	-0.002	-0.002	2.00E-04
	(0.007)	(0.006)	(0.002)	(0.002)	(0.006)	(0.005)	(0.002)	(0.002)
Treat	0.098**	0.048**	0.006*	0.014**	0.058**	0.031	0.014**	0.021**
	(0.016)	(0.01)	(0.003)	(0.004)	(0.013)	(0.024)	(0.005)	(0.006)
Black	-0.029**	-0.017*	0.006**	-0.0030	-0.032**	-0.010	0.010**	0.003
	(0.007)	(0.008)	(0.002)	(0.003)	(0.007)	(0.007)	(0.002)	(0.002)
Observations	26,543	18,443	26,543	18,443	34,788	20,608	34,788	20,608
Adjusted R2	0.018	0.015	0.002	0.006	0.021	0.009	0.008	0.003
Mean Y	0.186	0.107	0.013	0.018	0.141	0.092	0.029	0.014

This table reports the change in hit rate of vehicle searches and pedestrian frisks in Chicago and Philadelphia using the difference-in-difference specification in Equation 1. Panel A reports the impact of the protests while Panel B reports the impact of the pandemic. Hit rate is measured as the probability of finding either (1) any contraband (drugs or weapons) or (2) specifically firearms following a search/frisk. Data from 2016-2020 are used for Chicago and 2015-2020 for Philadelphia. For Pandemic regressions, observations range from 6 weeks before the COVID-19 pandemic onset to 4 weeks after. For Protest regressions, observations range from 3 weeks before the George Floyd Protests to 6 weeks after. The same calendar dates are used for all years. After = 1 beginning on the calendar day of the first day of the relevant event and 0 otherwise; Treat=1 for 2020 and 0 otherwise. Black=1 if the race of the driver/pedestrian was recorded by the officer to be Black. All regressions include police region and year fixed effects as well as controls for detainee age and gender (age is split into the following categories: <20,20-30,30-40,40-50,>50). Robust standard errors clustered at the region level. Data Source: Philadelphia Police Department & Chicago Police Department

Table 3A: Impact of the Protests Response on Stop Hit Rate

	Chicago				Philadelphia			
	All Contraband		Guns only		All Contraband		Guns only	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After*Treat	0.020	0.021**	0.002	0.011**	0.006	4.82E-04	0.003	0.005
	(0.013)	(0.007)	(0.003)	(0.004)	(0.005)	(0.009)	(0.002)	(0.004)
After	0.004	0.002	0.001	0.002*	-0.001	-3.99E-04	1.48E-04	1.06E-04
	(0.005)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0003)	(0.0003)
Treat	0.107**	0.032**	0.010**	0.018**	0.0313**	0.0291**	0.0049**	0.0055**
	(0.016)	(0.007)	(0.004)	(0.003)	(0.003)	(0.006)	(0.001)	(0.002)
Black	-0.016*	0.004*	0.003**	0.0010	0.0016*	0.0069**	0.0014**	0.0013*
	(0.007)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0003)	(0.001)
Observations	36,448	63,272	36,448	63,272	285,056	112,591	285,056	112,591
Adjusted R2	0.029	0.014	0.005	0.012	0.010	0.008	0.002	0.003
Mean Y	0.127	0.032	0.009	0.008	0.014	0.017	0.003	0.003

Table 3B: Impact of the Pandemic Response on Stop Hit Rate

	Chicago				Philadelphia			
	All Contraband		Guns only		All Contraband		Guns only	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After*Treat	0.014	0.020**	0.005	0.008**	0.0066*	0.010	0.0040**	0.002
	(0.012)	(0.006)	(0.003)	(0.003)	(0.003)	(0.006)	(0.001)	(0.002)
After	-0.003	-0.003	0.0001	-0.001	0.000	-0.001	-2.08E-04	-4.90E-05
	(0.005)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0002)	(0.0003)
Treat	0.071**	0.010*	0.004*	0.004**	0.0113**	0.0130**	0.0021**	0.0045**
	(0.011)	(0.004)	(0.001)	(0.001)	(0.002)	(0.005)	(0.001)	(0.001)
Black	-0.015**	0.004	0.003**	0.0002	0.0023*	0.0044**	0.0019**	0.0013**
	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0002)	(0.0004)
Observations	46,421	62,039	46,421	62,039	354,726	125,304	354,726	125,304
Adjusted R2	0.018	0.011	0.002	0.004	0.009	0.005	0.003	0.002
Mean Y	0.121	0.034	0.007	0.005	0.014	0.017	0.003	0.002

This table reports the change in hit rate of pedestrian frisks and vehicle stops using the difference-in-difference specification in Equation 1. Panel A reports the impact of the protests while Panel B reports the impact of the pandemic. Hit rate is measured as the probability of finding either (1) any contraband (drugs or weapons) or (2) specifically firearms following a stop. Data from 2016-2020 are used for Chicago and 2015-

2020 for Philadelphia. For Pandemic regressions, observations range from 6 weeks before the COVID-19 pandemic onset to 4 weeks after. For Protest regressions, observations range from 3 weeks before the George Floyd Protests to 6 weeks after. The same calendar dates are used for all years. After = 1 beginning on the calendar day of the first day of the relevant event and 0 otherwise; Treat=1 for 2020 and 0 otherwise. Black=1 if the race of the driver/pedestrian was recorded by the officer to be Black. All regressions include police region and year fixed effects as well as controls for detainee age and gender (age is split into the following categories: <20,20-30,30-40,40-50,>50). Robust standard errors clustered at the region level. Data Source: Philadelphia Police Department & Chicago Police Department

Table 4: Impact of Protests on Hit Rate Controlling for Active Officers

	Chicago				Philadelphia			
	All Contraband		Guns only		All Contraband		Guns only	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
After*Treat	0.045*	0.040	0.007	0.038**	0.036	-0.001	0.019	0.011
	(0.022)	(0.022)	(0.007)	(0.013)	(0.026)	(0.031)	(0.014)	(0.016)
After	-0.008	0.005	-2.00E-03	-0.004	0.002	-0.009	-2.00E-05	0.003
	(0.019)	(0.015)	(0.004)	(0.006)	(0.013)	(0.012)	(0.007)	(0.006)
Treat	-0.094**	-0.004	0.010	0.016	0.047	0.028	0.009	0.018
	(0.030)	(0.029)	(0.007)	(0.019)	(0.040)	(0.024)	(0.014)	(0.011)
Black	-0.005	-0.033	0.010*	0.011	-0.028	0.009	0.003	0.000
	(0.02)	(0.022)	(0.004)	(0.009)	(0.015)	(0.015)	(0.006)	(0.007)
Observations	7,254	4,256	7,254	4,256	7,309	3,187	7,309	3,187
Adjusted R^2	0.027	0.015	0.009	0.027	0.039	0.013	0.011	-0.001
Mean Y	0.253	0.142	0.020	0.049	0.172	0.109	0.034	0.021

Note: $^{*}p<0.05$ $^{**}p<0.01$

This table reports the change in hit rate of vehicle searches and pedestrian frisks conducted by officer pairs who conduct at least one stop in the 6 weeks after the George Floyd Protests began. The difference-in-difference specification in equation 1 is used. Hit rate is measured as the probability of finding either (a) any contraband (drugs or weapons) or (b) specifically firearms. Data from 2016-2020 for Chicago and 2015-2020 for Philadelphia are used. Observations range from 3 weeks before the start of the George Floyd Protests to 6 weeks after. The same calendar dates are used for all years. After = 1 beginning on 29 May if Chicago and 3 June in Philadelphia for each year and 0 otherwise; Treat=1 for 2020 and 0 otherwise. Black=1 if the race of the pedestrian was recorded by the officer to be Black. All regressions include sector and year fixed effects as well as controls for detainee age and gender (age is split into the following categories: <20,20-30,30-40,40-50,>50). Robust standard errors clustered at the sector level. Data Source: Chicago Police Department & Phiadelphia Police Department

Table 5: Legal Justification of Pedestrian Stops in Philadelphia

	Pedestrian Stops	
	Illegal Stop	Illegal Frisk
	(1)	(2)
After*Treat	-0.013	-0.065
	(0.031)	(0.099)
After	-0.014	-0.055
	(0.010)	(0.039)
Treat	-0.079**	-0.102
	(0.029)	(0.072)
Black	0.027	0.016
	(0.017)	(0.055)
Observations	5,771	812
Adjusted R^2	0.019	0.016
Mean Y	0.200	0.376
Note:		*p **p ***p<0.01

This table reports the change in the probability that a given pedestrian stop/frisk lacks legal justification. The difference-in-difference specification in equation 1 is used but with the dependent variable being a dummy for whether the stop/frisk was conducted illegally. Legality was determined by an audit of a randomly drawn sample of stops taken from the full set of pedestrian stops provided by the Philadelphia Police Department. Data from 2016-2020 is used. Observations range from 3 weeks before the start of the George Floyd Protests to 6 weeks after. The same calendar dates are used for all years. Illegality of a frisk is measured conditional on a frisk having occurred. After = 1 beginning on June 3 of each year and 0 otherwise; Treat=1 for 2020 and 0 otherwise. Black=1 if the race of the pedestrian was recorded by the officer to be Black. Robust standard errors clustered at the PSA level. All regressions include PSA and year fixed effects as well as controls for detainee age and gender (age is split into the following categories: <20,20-30,30-40,40-50,>50). Data Source: Philadelphia Police Department

Table 6: Change in Crime Following Protests

Panel A:	**Log of Daily Incident Reports (Chicago)**						
	Overall	Violent	Homicide	Rape	Shooting	Aggravated Assault	Robbery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
After*Treat	0.136**	0.186**	0.589**	-0.036	0.378*	0.232**	0.084
	(0.042)	(0.063)	(0.218)	(0.176)	(0.147)	(0.073)	(0.095)
After	0.046**	0.087**	-0.001	0.121	0.144**	0.061*	0.121**
	(0.009)	(0.023)	(0.085)	(0.069)	(0.053)	(0.030)	(0.028)
Treat	-0.396**	-0.239**	-0.252	-0.400**	-0.258*	-0.068	-0.548**
	(0.025)	(0.050)	(0.190)	(0.141)	(0.128)	(0.063)	(0.079)
Observations	315	315	315	315	315	315	315
Adjusted R^2	0.616	0.161	0.062	0.067	0.216	0.081	0.381

Panel B:	**Log of Daily Incident Reports (Philadelphia)**						
	Overall	Violent	Homicide	Rape	Shooting	Aggravated Assault	Robbery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
After*Treat	-0.287**	-0.052	-0.118	0.588**	-0.107	-0.035	-0.266**
	(0.056)	(0.063)	(0.239)	(0.221)	(0.182)	(0.074)	(0.101)
After	0.006	-0.005	0.068	-0.021	0.072	-0.004	-0.005
	(0.013)	(0.021)	(0.088)	(0.077)	(0.075)	(0.030)	(0.035)
Treat	-0.144**	-0.160**	0.335	-0.989**	0.565**	0.11	-0.529**
	(0.052)	(0.056)	(0.223)	(0.196)	(0.169)	(0.065)	(0.081)
Observations	378	378	378	378	378	378	378
Adjusted R^2	0.444	0.094	0.012	0.088	0.079	0.017	0.339

This table reports the change in daily crime from the start of the protests using the difference-in-difference specification in equation 1 with crime reports as the dependent variable. For homicide, rape and shooting regressions, 0.5 is added to the daily number to account for days with zero incidents. Each column reports a separate regression. Panel A reports the results for Chicago, Panel B for Philadelphia. Data from 2016-2020 are used for Chicago and 2015-2020 for Philadelphia.. Observations range from 3 weeks before the beginning of the George Floyd Protests and 6 weeks after for each year. After = 1 beginning June 3 for Philadelphia, May 29 for Chicago and is 0 otherwise; Treat = 1 for 2020 and 0 otherwise. All regressions include year fixed effects. Robust standard errors reported. Data Source: Chicago Police Department and Philadelphia Police Department

Table 7: Changes in Selected Subset of Chicago

Panel A:	Log of Daily Incident Reports						
	Overall	Violent	Homicide	Rape	Shooting	Aggravated Assault	Robbery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
After*Treat	0.103*	0.108	0.271	-0.029	0.056	0.198*	-0.138
	(0.046)	(0.068)	(0.225)	(0.211)	(0.161)	(0.081)	(0.112)
After	0.024*	0.065**	0.105	0.165	0.240**	0.026	0.136**
	(0.011)	(0.024)	(0.097)	(0.102)	(0.073)	(0.033)	(0.040)
Treat	-0.288**	-0.123*	-0.032	-0.176	-0.045	-0.0005	-0.352**
	(0.033)	(0.054)	(0.202)	(0.186)	(0.138)	(0.068)	(0.093)
Observations	315	315	315	315	315	315	315
Adjusted R^2	0.394	0.074	0.039	0.009	0.161	0.050	0.222

	Protests				Pandemic			
	All Contraband		Guns only		All Contraband		Guns only	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After*Treat	0.061**	0.061*	0.004	0.040	0.003	0.032	0.002	0.034**
	(0.023)	(0.028)	(0.009)	(0.021)	(0.02)	(0.023)	(0.006)	(0.013)
After	0.002	0.013	-0.001	0.013**	0.003	-0.023**	0.002	-5.00E-03
	(0.009)	(0.008)	(0.003)	(0.004)	(0.007)	(0.007)	(0.002)	(0.003)
Treat	0.154**	0.087**	0.017*	0.073**	0.095**	0.040**	0.005	0.013*
	(0.018)	(0.022)	(0.007)	(0.015)	(0.012)	(0.014)	(0.004)	(0.006)
Black	-0.033*	-0.029	0.008*	-0.0110	-0.023	-0.025	0.008*	0.004
	(0.015)	(0.017)	(0.004)	(0.009)	(0.013)	(0.015)	(0.003)	(0.006)
Observations	10,885	8,361	10,885	8,361	14,593	9,481	14,593	9,481
Adjusted R^2	0.026	0.017	0.004	0.027	0.014	0.012	0.002	0.006
Mean Y	0.207	0.127	0.020	0.036	0.187	0.119	0.015	0.020

Note: $^{*}p<0.05$ $^{**}p<0.01$

This table shows results for the subset of Chicago police sectors highlighted in figure 7. Panel A show the change in daily crime from the start of the protests using the difference-in-difference specification in equation 1 with crime reports as the dependent variable. For homicide, rape and shooting regressions, 0.5 is added to the daily number to account for days with zero incidents. Panel B reports the change in the hit rate of vehicle searches and pedestrian frisks. Each column reports a separate regression. Data from 2016-2020 are used for the 8 police districts which were had some of the lowest increases in shootings over this period. Observations range from 3 weeks before the beginning of the George Floyd Protests and 6 weeks after for each year. After = 1 beginning May 29 and 0 otherwise; Treat = 1 for 2020 and 0 otherwise. All regressions include year fixed effects. Robust standard errors reported. Data Source: Chicago Police Department

Table 8: Impact of Protests on Police Hit Rate by Race

	Chicago				Philadelphia			
	All Contraband		Guns only		All Contraband		Guns only	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
After*Treat*Black	-0.003	-0.004	0.003	-0.018	-0.039	0.045	-0.030	-1.00E-04
	(0.035)	(0.037)	(0.01)	(0.023)	(0.042)	(0.068)	(0.029)	(0.031)
After	0.010	0.013	0.003	0.004	-0.023	-0.017	-0.005	-0.002
	(0.012)	(0.01)	(0.002)	(0.005)	(0.015)	(0.016)	(0.006)	(0.005)
Treat	0.071**	0.043*	0.006	0.010	0.042	0.031	0.005	0.009
	(0.02)	(0.02)	(0.005)	(0.009)	(0.028)	(0.035)	(0.017)	(0.013)
Black	-0.027*	-0.011	0.006*	-0.010*	-0.051**	-0.024	0.001	-2.00E-04
	(0.012)	(0.01)	(0.003)	(0.005)	(0.015)	(0.012)	(0.006)	(0.006)
After*Treat	0.033	0.033	0.001	0.036*	0.075*	-0.050	0.040	0.014
	(0.029)	(0.029)	(0.007)	(0.017)	(0.029)	(0.064)	(0.026)	(0.025)
After*Black	-0.006	0.003	-0.003	0.005	0.022	0.024	0.010	0.004
	(0.014)	(0.012)	(0.003)	(0.006)	(0.015)	(0.018)	(0.007)	(0.006)
Treat*Black	0.064**	0.033	0.012	0.058**	0.037	0.004	0.008	0.011
	(0.024)	(0.025)	(0.007)	(0.015)	(0.027)	(0.040)	(0.018)	(0.015)
Observations	20,024	16,794	20,024	16,794	25,507	15,041	25,507	15,041
Adjusted R2	0.020	0.014	0.006	0.018	0.024	0.006	0.006	0.006
Mean Y	0.195	0.112	0.016	0.029	0.148	0.108	0.031	0.018

Note: $^{*}p<0.05$ $^{**}p<0.01$

This table reports the change in hit rate of pedestrian frisks and vehicle searches using the difference-difference-in-difference specification in Equation 3. Hit rate is measured as the probability of finding either (1) any contraband (drugs or weapons) or (2) specifically firearms. Data from 2016-2020 for Chicago and 2015-2020 for Philadelphia are used.. Observations range from 3 weeks before the George Floyd Protests to 6 weeks after. The same calendar dates are used for all years. After = 1 beginning on May 29 each year and 0 otherwise; Treat=1 for 2020 and 0 otherwise. Black=1 if the race of the pedestrian/driver was recorded by the officer to be Black. All regressions include year and police region fixed effects as well as controls for detainee age and gender (age is split into the following categories: <20,20-30,30-40,40-50,>50). Robust standard errors clustered at the region level. Data Source: Chicago Police Department & Philadelphia Police Department.

Appendix

Table A1: Impact of Protests on the Number of Stops and Searches/Frisks

Panel A: Chicago

	Log of Daily Stops		Log of Daily Searches/Frisks		Log of Daily Stops (Active Officers)		Log of Daily Searches/Frisks (Active Officers)	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After*Treat	-0.568**	-0.606**	-0.586**	-0.444**				
	(0.082)	(0.085)	(0.085)	(0.087)				
After	-0.085**	0.037	-0.122**	-0.066*	-0.653**	-0.568**	-0.708**	-0.510**
	(0.029)	(0.025)	(0.033)	(0.029)	(0.077)	(0.082)	(0.078)	(0.082)
Treat	0.401**	-0.517**	0.517**	-0.292**				
	(0.056)	(0.073)	(0.057)	(0.071)				
Observations	315	315	315	315	63	63	63	63
Adjusted R^2	0.680	0.832	0.579	0.642	0.444	0.442	0.450	0.346
Note:								*p<0.05 **p<0.01

Panel B: Philadelphia

	Log of Daily Stops		Log of Daily Searches/Frisks		Log of Daily Stops (Active Officers)		Log of Daily Searches/Frisks (Active Officers)	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After*Treat	-1.345**	-1.457**	-1.349**	-1.690**				
	(0.083)	(0.170)	(0.103)	(0.157)				
After	-0.039	-0.039*	-0.085*	-0.070	-1.497**	-1.384**	-1.760**	-1.434**
	(0.025)	(0.020)	(0.035)	(0.038)	(0.170)	(0.079)	(0.153)	(0.098)
Treat	-1.890**	-0.551**	-1.184**	0.529**				
	(0.056)	(0.161)	(0.077)	(0.127)				
Observations	378	378	378	378	63	63	63	63
Adjusted R^2	0.929	0.869	0.828	0.616	0.648	0.772	0.661	0.719
Note:								*p<0.05 **p<0.01

This table reports the change in the number of stops, vehicle searches and pedestrian frisks following the Protests. Panel A reports the results for Chicago, Panel B for Philadelphia Models (1) and (2) report the change in daily vehicle and pedestrian stops respectively. Models (3) and (4) report the change in daily vehicle searches and pedestrian frisks respectively. Models (5) and (6) report the change in daily stops using only stops conducted by officer pairs who conduct at least one stop in the 6 weeks after the George Floyd Protests began. Models (7) and (8) report the change in daily vehicle searches and pedestrian frisks using this same set of officers. Models (1)-(4) use data from 2016-2020 for Chicago and 2015-2020 for Philadelphia while Models (5)-(8) use only 2020 data. Observations range from 3 weeks before the start of the George Floyd Protests to 6 weeks after. The same calendar dates are used for all years. After = 1 beginning on June 3 and 0 otherwise. Robust standard errors clustered at the police region level. Models (1)-(4) include year fixed effects. Data Source: Chicago Police Department & Philadelphia Police Department

Table A2: Impact of Protests on Search/Frisk Hit Rate, 2020 Only

	Chicago				**Philadelphia**			
	All Contraband		Guns only		All Contraband		Guns only	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
After	0.032*	0.043*	0.005	0.036**	0.036*	-0.002	0.020*	0.014
	(0.015)	(0.018)	(0.005)	(0.013)	(0.017)	(0.031)	(0.009)	(0.024)
Black	-0.001	-0.027	0.018*	0.026	0.008	0.006	0.006	-0.003
	(0.021)	(0.026)	(0.007)	(0.017)	(0.019)	(0.034)	(0.010)	(0.018)
Observations	3,613	1,941	3,613	1,941	3,010	801	3,010	801
Adjusted R2	0.035	0.030	0.008	0.024	0.048	0.004	0.020	-0.037
Mean Y	0.274	0.170	0.028	0.076	0.220	0.136	0.042	0.034

This table reports the change in hit rate of vehicle and pedestrian searches/frisks using a single difference (before-after) specification. Hit rate is measured as the probability of finding either (a) any contraband (drugs or weapons) or (b) specifically firearms following a search/frisk. Data from 2020 are used with observations ranging from 3 weeks before the start of the George Floyd Protests to 6 weeks after. After = 1 beginning on 3 June 2020 and 0 otherwise; Black=1 if the race of the driver/pedestrian was recorded by the officer to be Black. All regressions include police region fixed effects as well as controls for detainee age and gender (age is split into the following categories: <20,20-30,30-40,40-50,>50). Robust standard errors clustered at the region level. Data Source: Chicago Police Department & Philadelphia Police Department

Table A3: Change in Composition of Motor Vehicle Crash Patients

	Age	
	Chicago	Philadelphia
Chi-Square		
$\chi 2$	8.303	8.098
df	4	4
p-value	0.081	0.088

This table reports the results of a Chi-square test on the distribution of age of motor vehicle crash patients before and after the Protests. Observations range from 3 weeks before the start of the George Floyd Protests to 6 weeks after. For the Chi-square test, age is split into the following categories: <20,20-30,30-40,40-50,>50. Data Source: Chicago Open Data and Philadelphia Trauma Outcomes Study (PTOS)

Table A4: Robustness Tests

Panel A: Chicago

	12 Week After Period				Excluding First Week after Protests Began				Excluding First Two Weeks after Protests Began			
	All Contraband		Guns only		All Contraband		Guns only		All Contraband		Guns only	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
After*Treat	0.044**	0.031*	-0.001	0.016	0.045**	0.030	0.001	0.026*	0.043**	0.042*	0.002	0.029*
	(0.014)	(0.015)	(0.005)	(0.01)	(0.016)	(0.018)	(0.006)	(0.012)	(0.016)	(0.018)	(0.006)	(0.013)
After	0.010	0.020**	0.001	0.009**	0.008	0.016**	0.001	0.009**	0.009	0.018**	0.001	0.010**
	(0.006)	(0.005)	(0.002)	(0.002)	(0.006)	(0.005)	(0.002)	(0.003)	(0.006)	(0.005)	(0.002)	(0.003)
Treat	0.118**	0.072**	0.015**	0.052**	0.113**	0.068**	0.014**	0.049**	0.116**	0.068**	0.013**	0.050**
	(0.012)	(0.013)	(0.004)	(0.008)	(0.013)	(0.013)	(0.004)	(0.009)	(0.013)	(0.014)	(0.004)	(0.009)
Black	-0.019**	-0.011	0.008**	0.003	-0.018*	-0.005	0.009**	0.000	-0.016	-0.011	0.008**	-0.001
	(0.007)	(0.006)	(0.002)	(0.003)	(0.008)	(0.008)	(0.002)	(0.004)	(0.009)	(0.008)	(0.002)	(0.004)
Observations	33,253	28,649	33,253	28,649	19,567	16,324	19,567	16,324	19,451	16,138	19,451	16,138
Adjusted R^2	0.023	0.016	0.006	0.016	0.019	0.014	0.006	0.017	0.020	0.016	0.006	0.018
Mean Y	0.203	0.120	0.016	0.030	0.198	0.113	0.016	0.029	0.199	0.115	0.016	0.030

Panel B: Philadelphia

	12 Week After Period				Excluding First Week after Protests Began				Excluding First Two Weeks after Protests Began			
	All Contraband		Guns only		All Contraband		Guns only		All Contraband		Guns only	
	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian	Vehicle	Pedestrian
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
After*Treat	0.035**	-0.006	0.034**	0.027*	0.062**	-0.020	0.027**	0.016	0.053**	-0.018	0.040**	0.014
	(0.013)	(0.02)	(0.007)	(0.011)	(0.019)	(0.027)	(0.01)	(0.016)	(0.018)	(0.027)	(0.011)	(0.016)
After	-0.003	0.000	0.005*	0.001	-0.007	0.005	0.002	0.003	-0.009	0.004	0.002	0.003
	(0.004)	(0.005)	(0.002)	(0.002)	(0.005)	(0.005)	(0.002)	(0.002)	(0.005)	(0.005)	(0.002)	(0.002)
Treat	0.065**	0.036*	0.014**	0.017*	0.075**	0.040*	0.014**	0.023**	0.075**	0.040*	0.012*	0.023**
	(0.01)	(0.016)	(0.005)	(0.008)	(0.011)	(0.016)	(0.005)	(0.008)	(0.011)	(0.016)	(0.005)	(0.008)
Black	-0.028**	-0.007	0.012**	0.002	-0.033**	-0.009	0.008**	0.004	-0.033**	-0.011	0.010**	0.002
	(0.005)	(0.006)	(0.002)	(0.002)	(0.007)	(0.008)	(0.003)	(0.003)	(0.007)	(0.008)	(0.003)	(0.003)
Observations	43,621	25,539	43,621	25,539	25,487	14,917	25,487	14,917	25,438	14,849	25,438	14,849
Adjusted R^2	0.021	0.007	0.007	0.006	0.024	0.008	0.006	0.006	0.023	0.010	0.007	0.007
Mean Y	0.149	0.107	0.034	0.019	0.151	0.107	0.032	0.018	0.151	0.107	0.033	0.018

Note: $^{*}p<0.05$ $^{**}p<0.01$

This table reports the change in hit rate of vehicle and pedestrian searches using the difference-in-difference specification in equation 1. Panel A reports the results for Chicago, Panel B for Philadelphia. Hit rate is measured as the probability of finding either (a) any contraband (drugs or weapons) or (b) specifically firearms. For Chicago: In all models, observations start from the 3 weeks before May 29. In models (1)-(4) 12 weeks after May 29 are used, in models (5)-(8), 6 weeks after 4 June are used with the days of protest in between excluded, in models (9)-(12), 6 weeks after 11 June are used with the days of protest in between excluded. For Philadelphia: In models (1)-(4), observations start from the 3 weeks before June 3, while in all other specifications, observations start from 3 weeks before the start of the George Floyd Protests on May 30. In models (1)-(4) 12 weeks after are June 3 are used, in models (5)-(8), 6 weeks after 5 June are used with the days of protest in between excluded, in models (9)-(12), 6 weeks after 12 June are used with the days of protest in between excluded. The same calendar dates are used for each year (2015-2020). . After = 1 beginning on 29 May for Chicago and 3 June for Philadelphia and 0 otherwise; Treat=1 for 2020 and 0 otherwise. Black=1 if the race of the driver/pedestrian was recorded by the officer to be Black. All regressions include sector and year fixed effects as well as controls for detainee age and gender (age is split into the following categories: <20,20-30,30-40,40-50,>50). Robust standard errors clustered at the sector level. Data Source: Chicago Police Department and Philadelphia Police Department