Do Cops Know Who to Stop?

Assessing Optimizing Models of Police Behavior with a Natural Experiment[*]

David S. Abrams, Hanming Fang, and Priyanka Goonetilleke

February 6, 2025

ABSTRACT

We investigate police stops using two unique events in 2020, the COVID-19 pandemic and George Floyd protests. Consistent with optimizing models of policing, contraband hit rates generally rose as stops fell; we rule out numerous alternative explanations, including changes in street population, crime, police allocation, and effort. We produce the first "policing production functions" and estimate elasticities. There is substantial variation by location and stop type, with some stops no more productive than random. We find no evidence that police stops decrease crime, at least in the short run, and mixed evidence about changes in racial disparities.

JEL Codes: K00, K14, J15

Keywords: Policing Reform; Stop and Frisk; Outcome Test

---

[*] We would like to thank David Bjerk, Aaron Chalfin, Mitch Downey, Jens Ludwig, Conrad Miller, Steve Raphael, Steve Ross, and participants in numerous seminars and conferences for useful comments and suggestions. We would like to acknowledge the generous financial support from the Data-Driven Discovery Fund at the University of Pennsylvania. All remaining errors are our own.

Abrams: Carey Law School and Wharton Business Economics and Public Policy, University of Pennsylvania, Philadelphia, PA 19104. Email: dabrams@upenn.edu; Fang: Department of Economics, University of Pennsylvania, Philadelphia, PA 19104. Email: hanming.fang@econ.upenn.edu; Goonetilleke: Pritzker School of Law, Northwestern University, Chicago, Il 60611. Email: priyanka.goonetilleke@law.northwestern.edu.

I. Introduction

Policing plays an important role in solving crime, deterring potential crime, and providing protection to the public in general. Recently, police-citizen interactions have come under renewed scrutiny. Following the death of George Floyd in 2020, there were calls to "defund the police," a slogan that supports divesting funds from police departments and reallocating them to non-policing forms of public safety and community support, such as social services, youth services, housing, education, healthcare and other community resources. Defunding the police, however, does not come without costs in the form of potentially increasing crime, due to lack of deterrence and punishment.[1]

However, not all forms of police activity are necessarily equally effective in deterring crime. Theoretically, the extent to which policing activity deters crime depends on its effectiveness in catching criminals. The more effective police activity is, the greater its effect on a criminal's probability of punishment and thus the greater its deterrent effect. Therefore, the key to responsible police reform is an understanding of the effectiveness of police activities.

Our paper is motivated by a desire to contribute to the policing reform debate by empirically estimating the relationship between contraband detection rate and measures of policing activity in the context of the "stop and frisk" programs in Chicago and Philadelphia, two of the largest police departments in the United States.

In addressing the important policy-relevant empirical question regarding the returns to "stop and frisk" activities in terms of contraband discoveries, we also contribute to the testing of an essential prediction of the optimizing models of police behavior. Becker (1968) deserves much of the credit for broadening economics to include the study of crime and race, and for suggesting the use of outcome tests. But Knowles, Persico, & Todd (2001), henceforth 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. The hit rate is simply the fraction of stops or searches that yields a weapon or some other type of contraband. An implication of KPT and almost all subsequent models on the subject is of diminishing marginal returns to stops (see Section II for more

[1] The deterrent effect of police is one of the most robust empirical findings in the law and economics of crime literature. See Chalfin and McCrary (2017) for an excellent review.

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.[2] But 2020 was a rare year, and brought not one, but two events where the number of 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 protests. We find that hit rates generally increased as police stops plunged, consistent with 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, the composition of individuals on the street, changes in police deployment, or crime. While both the total number of police making stops and the number of stops per officer 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.

We define the policing production function as the relationship between hit rates and frisk rates and compute elasticities for different combinations of frisk type and contraband. We find that when contraband is defined as firearms alone, elasticities are always higher than when defined as all contraband (including drugs), which suggest that the police respond more to information predicting firearm possession versus other types of contraband. We also find that with one exception, elasticities are higher for pedestrian than the analogous vehicle stops, which suggests that officers have more predictive information from individuals on foot, rather than in vehicles. These results complement those presented in Feigenbaum & Miller (2021) which found elasticities close to zero for Texas highway stops, using between-officer comparisons. Here we are able to use within-officer comparisons for both pedestrian and vehicular stops in major cities.

[2] In a separate context, the issuance of traffic fines, Goncalves and Mello (2023) find that officers do not apply discretion in a manner consistent with minimizing future traffic offending and crash involvement.

KPT (2001) initiated a large literature in economics, reviewed below in Section II.C, that studies racial bias in all aspects of law enforcement and the judicial system. Aspects studied include policing, highway patrols, bail setting, sentencing, probation, and 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 it has elicited a great deal of theoretical and empirical scrutiny. Although many aspects of KPT have been challenged, almost all subsequent research models officers as agents seeking to maximize an objective function related to stops or searches. This payoff typically includes a weighted average of a legally justified component e.g. finding contraband, and an illegitimate component, e.g. racial prejudice.

We use the exogenous changes in policing due to the pandemic and protests to investigate the predictions of such optimizing policing models.[3] Our estimates shed light on parameters that are important for making policy decisions about policing. While the changes in police stops were large for both events, frisks were almost flat during the onset of the Covid pandemic and mobility declined sharply at that time, causing potential confounds. Thus, most of our analysis is focused on the protests, when mobility changes were small, but frisks plummeted alongside stops.

We perform several additional tests of the protest period in order to rule out the possibility that our results are due to contemporaneous changes that may also affect the hit rates. In order to test possible changes in traffic composition, we obtain hospital data including the age distribution of individuals involved in accidents, and we find no significant changes in the period. Even though the numbers and some measure of the composition of the relevant population is likely constant in the time window of our study, it is still possible that the share of potential criminals grew. However, examining crime data we find that in this period crime rates in Philadelphia are mostly either flat or decreasing, depending on the crime categories. While

[3] Abrams, Fang and Goonetilleke (2022) conducts a more limited analysis exclusively on the effect of the protests in Chicago.

violent crime did rise in Chicago during this period, focusing on a region of the city where violent crime did not rise substantially, we still find substantial increases in hit rates.

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 of the frisks.

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.[4] 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.

As argued, 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 policy implications from our results are nuanced: on the one hand, police officers are effective in finding contraband using "stop and frisk;" on the other hand, such policing practices have little deterrence effect on crime.

The rest of the paper proceeds as follows. Section II provides a brief legal background on police stops, describes the simple theoretical intuition for our tests, and reviews the most relevant literature; Section III introduces the data on police stops in Chicago and Philadelphia; Section IV presents the main results; Section V adds robustness checks with a discussion of potential threats to validity; finally, Section VI concludes.

[4] Bacher-Hicks and de la Campa (2020) similarly do not find evidence that stop and frisk reduces overall criminal incidents.

II. Background

A. Legal Background

The focus of this paper is policing in Chicago and 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 a useful deterrent against crime, although we are aware of no well-identified study to this effect. [5] 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 only if there is probable cause – a higher legal standard - 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* and *Floyd v. City of New York*, both of which led to settlement agreements that have been in force for over a decade. [6]

B. Testable Implications of Optimizing Model of Policing

Most of the literature on optimizing police behavior relies on Becker's outcome test, which in this setting can be described as follows. Assume that the objective of police officers described by these models is to detect the presence of contraband on drivers, passengers or pedestrians.[7] If a police officer is racially prejudiced against a minority group, he or she would choose to stop and/or search a minority group member with less convincing evidence for the presence of

---

[5] But see e.g. MacDonald and Braga (2019) for excellent exploratory analysis of changes in New York City's stop-and-frisk program after litigation.

[6] See Abrams (2014) for further background about stop-and-frisk policing from a law and economics perspective. Gelman, Fagan and Kiss (2007) examines stop rates in NYC controlling for observables.

[7] Contraband is primarily weapons but in Section III we discuss broadening it to include drugs and other objects.

contraband. Thus, in the aggregate, the contraband discovery rates among the stopped/searched minority group members should be lower than non-minorities. Therefore, a comparison of outcomes across different groups can be suggestive of the presence of racial prejudice by the police.

However, as argued in the literature, the *infra-marginality* problem may complicate the application 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 an individual if the suspicion level is above a threshold; an officer prejudiced against non-White people will use a lower suspicion threshold for members of that group than for White individuals. Thus, theoretically the contraband finding rate among the *marginal* minority individual and that among the *marginal* White individual are indicative of the officer's prejudice. However, the outcome tests in the literature often compare the *average* contraband finding rates against different groups. The comparison of the group averages may *not* be in the same direction as the comparison of the marginals.

Knowles, Persico, & Todd (2001) 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. 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 the contraband detection rate, at least as part of their objective function, will search drivers only if their suspicion for the driver carrying contraband exceeds a threshold. That is, they will allocate search effort only to those deemed more likely to carry contraband.

If officers are prejudiced against certain groups of drivers, then they will use 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 comparing officers of different races and use officers of a given race as a benchmark 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, both models assume that the police officers are rational and are trying to optimize an objective that includes finding contraband as one of its components. As such, both models would predict that, if for some exogenous reason the costs of stopping or searching drivers (or pedestrians) were to go up -- as they would during a pandemic or amidst a nationwide protest against police brutality -- the contraband finding rates against all drivers should go up.

In the 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 officer's search cost 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, *ceteris paribus.*

To illustrate these ideas more clearly, let us consider a very simple optimizing model of police behavior along the lines of Anwar and Fang (2006). Suppose that there is a continuum of pedestrians (or drivers), who may or may not be carrying weapons or other contraband. Police officers may not perfectly observe whether the pedestrian is guilty, but based on the contextual signals, including e.g., demeanor, time of the day, location, etc., they form a belief (or suspicion) regarding the likelihood that the pedestrian is carrying contraband or weapons. Let us denote this suspicion by $\theta \in [0,1]$ where a θ close to 1 indicates high confidence by the officer that the subject is guilty. In the population of pedestrians (or drivers), the suspicion θ is distributed according to density function $p(\theta)$.

In an optimizing model of police behavior, an officer assesses for each subject whether that individual should be frisked, based on the benefit and cost of frisking the subject. The benefit from frisking the subject is assumed to arise from successfully recovering contraband or a weapon from the subject: if an officer frisks a pedestrian and finds contraband (a "hit"), his/her payoff is normalized to 1; otherwise the payoff is 0; the total cost of frisking includes the effort cost needed to execute the frisk, denoted by $C \in [0,1]$ and a shadow cost of the time spent on the frisk, denoted by $\lambda \in [0,1]$, which may be affected by the search capacity of the police, for example. Formally, the officer will frisk a subject of suspicion θ if

$$\theta > C + \lambda \equiv \theta^*$$

In other words, an optimizing police officer will selectively target their costly frisks to the subjects whose suspicion level exceeds the threshold θ^*. If for some exogenous reasons, the police now face a lower frisking capacity, the shadow cost of search increases from λ to $\lambda' > \lambda$. The officer will then increase his/her search threshold to $\theta^{*\prime} > \theta^*$ (See Figure 1 below). Thus under an optimizing model of police search behavior, we have the robust prediction that an exogenous reduction in frisk capacity will increase the average hit rate from fewer frisks, as the optimizing police officer will reserve his/her fewer frisks to the most suspicious subjects (i.e., those subjects whose probability of carrying contraband exceeds $\theta^{*\prime}$, which is higher than the threshold θ^* used prior to the reduction in search capacity).

Figure 2 shows the policing production function under two assumptions, random and targeted policing. Here we assume hit rate is the single measure of productivity and the input is stops or frisks. If police do not know whom to stop, i.e., if the police are simply randomly stopping pedestrians and/or drivers, then an exogenous reduction in policing activities would linearly lower the total number of contraband findings (Panel A of Figure 2, left); equivalently, the contraband finding rate will be *constant* with respect to the policing activities (Panel B of Figure 2, left). By contrast, if police are effective in targeting -- probabilistically at least-- those with contraband, then there will be a concave relationship between the total number of contraband findings and the policing activities (Panel A of Figure 2, right); equivalently, a reduction in policing activities will increase the contraband finding rate (Panel B of Figure 2, right).

The intuition for the above differences between random vs. targeted policing is very straightforward. If police are simply randomly stopping/frisking pedestrians/drivers, then the contraband finding rate will be equal to the underlying contraband carrying rate of the population, independent of how many are stopped/frisked. By contrast, if police engage in targeted policing, the individuals who are judged to be the most suspicious according to the information/signals deemed relevant by the police in predicting criminality will be stopped/frisked first. Only as the number of stops/frisks increase will the less suspicious pedestrians/drivers be stopped. Thus, the contraband finding rate will decrease in the total number of stops/frisks, and the total number of contraband "hits" will exhibit diminishing returns with respect to the number of stops/frisks.

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. All that is needed is an exogenous shock to police activity. We describe two such plausible shocks occurring in the year 2020 in Section III.

C. Related Literature

Dharmapala and Ross (2004), Anwar and Fang (2006), and Antonovics & Knight (2009) discuss 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. Anwar and Fang (2006) and Antonovics and Knight (2007) argued that KPT's test may not be robust when its model is generalized to allow for trooper heterogeneity.

Two recent papers have sought to empirically test the underlying assumption that police officers face diminishing returns to searches. Feigenberg and Miller (2022) 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, i.e., that the conditional hit rate is roughly constant across troopers with different search rates.

Meanwhile, Gelbach (2021) further elaborates that a testable implication of an optimizing model of officer search behavior of Anwar and Fang (2006) is that at the *officer level* there should be a negative relationship between search rates and conditional hit rates. He empirically tests this implication for Florida and Harris County Texas and 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.[8]

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 (2022) 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.

The above literature assumes that the decision maker, e.g., the officer or the pre-trial judges, holds correct beliefs regarding the underlying distribution of outcomes of concern among the treated – whether the drivers carry contraband for the officers and whether the defendant will fail to appear to court for the pre-trial judges. A recent literature started to incorporate *incorrect* statistical beliefs into the analysis. Bohren et al (2020) shows that distinguishing between racial bias due to racial animus and biased beliefs is challenging, as preferences and beliefs can manifest equivalently in a judge's decisions. However, prejudice and incorrect statistical beliefs can be distinguished if additional sources of variations are present in the data. In Bohren et al (2020), the source of variation is informational provision, which could change beliefs not animus. Hull (2021) argues that the slope of the marginal treatment effect curve can be used to reject accurate statistical discrimination and canonical taste-based discrimination. In our analysis, the source of variation we exploit is the within-officer search capacity; and the predictions hold

[8] Gelbach (2021) also tested the aforementioned negative relationship for the pre-trial bail setting, as analysed in Arnold, Dobie and Yang (2018, ADY), where the prediction for an optimizing judge is that the pre-trial rearrest rates and the pre-trial release rates should be negatively correlated. He found that ADY data violates this specification test.

both when the officers have rational and incorrect beliefs about the underlying distribution of criminality in the population.

III. Data

We examine police stops in Chicago and Philadelphia and thus our two main data sets come from those cities' police departments. The Philadelphia data set contains suspect-stop level data about all pedestrian and vehicle police stops within the city from January 1, 2015 to December 31, 2020.[9] After we remove the 2 observations for which location data is incomplete, the data set comprises 2,794,918 stops. The data includes information on the timing and the location of the stop, ID of the police officer conducting the stop, demographic information on the individual stopped,[10] whether a frisk, search or arrest occurred, vehicle information, any contraband discovered, and a narrative field. The other primary data set contains similar variables for Chicago, spanning from January 1, 2016 to December 31, 2020, and comprises 564,935 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 3a and 3b, which display an index of pedestrian stops for 2020 (solid black) and prior years (dashed gray) for Chicago and Philadelphia, respectively. The figures are normalized so that the mean stop level from January 1 to January 7 of each series equals 100. The solid red vertical line indicates the start of the week in which the Philadelphia Police Department announced a new policy for nonviolent incidents while the dashed green line marks the start of the decline in police stops following the protests and looting in response to the George Floyd killing. In mid-March 2020, there was a 61% and 33% decline in pedestrian police stops in Chicago and in Philadelphia respectively**.** After increasing back towards normal levels, there was again a dramatic decline of 44% in

---

[9] 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.

[10] Demographic data is reported by the officer conducting the stop and is thus subject to potential misreporting (Luh 2022). However, as this demographic data is not used in our main specification, any potential misreporting would not impact our primary analysis.

Chicago and 72% in Philadelphia in stops following the killing of George Floyd.[11] The policing changes around these two unanticipated events are the focus of the paper.

The bottom panels of Figures 3a and 3b show analogous changes in pedestrians frisks. In Chicago the drop in the number of 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 again, by 39%. In Philadelphia, the pattern for frisks is somewhat different for the first event, with a sharp decline followed by a swift return to prior levels of frisks. But the change in frisks in the period around the protests is very similar to stops with a precipitous decline of 73%, and a new lower level for frisks that stays relatively constant through most of the summer.

Figure 4 displays 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 51% and 79% 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 29%. The decline following the onset of the protests was 48% in Chicago and 81% 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 constant the set of officers (columns 5-8). The figures establish the sudden changes in policing during these two moments in 2020, but the key to use these changes as a source of variation to test for optimizing models of policing behavior is to tease out whether there were other relevant concurrent changes.

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. [12] While the number of stops declines sharply after both events, it is worth examining changes in their characteristics separately for each event.

---

[11] We refer to the dramatic decline in police stops surrounding the pandemic and the protests as the two "events" in this study.

[12] 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.

For the pandemic event, the share of stops leading to a frisk increased substantially in both cities. For pedestrian stops the frisk rate increased 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 Philadelphia and 5.9 percentage points in Chicago. 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 and Newall, 2020). In order to reduce the number of people 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 of people stopped was relatively unchanged in Chicago around the pandemic onset, in Philadelphia the Black share increased by around 10% for both pedestrian and vehicle stops. The mean age of detainees from vehicle stops dropped by more than 1.5 years 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 (1.8 percentage points for contraband and 2.4 percentage points for guns).

We now turn to the policing change around the protests. In Chicago the change in the Black share of pedestrian stops was again quite small, however there was a 10% increase in the Black share of vehicle stops. In comparison, in Philadelphia the change in the 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. Thus, the drop in the number of frisks and searches conducted appears to be driven by a drop in the rate of stops. Under an optimizing model of behavior, if officers observe information before stopping an individual and update this information following the stop, then a constant search/frisk rate combined with an exogenous decrease in searches implies a higher threshold for conducting a search/frisk. This is because the average suspicion of stopped individuals is higher following the decrease in stops.

Following the protests, hit rates for guns 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 (1.1 percentage points), these also generally rise. This drop in the contraband hit rate in Philadelphia was driven by a large fall of 4.6 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.

We make use of two additional types of data to examine the potential changes in the population at risk of being stopped. The first contains information about the driving population using data from automobile crashes. For Chicago this data also comes from the Chicago Police Department. The 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. It includes time and location of the crash as well as demographic information of the individuals involved. For Philadelphia we obtain data 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 aims to estimate the impact of a large change in police frisks on the share of frisks that result in contraband discovery (the frisk "hit rate"). Under a model of optimizing behavior, officers maximize the likelihood of contraband detection from the marginal frisk. Hence, a large decline in frisks should substantially increase the hit rate. To test for this, 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:

$$contraband_i = \alpha + \beta_1 after_i * treat_i + \beta_2 after_i + \sum_{k=y_{min}}^{2020} (\gamma_k * year_{ik}) + \sum_{j=1}^{J} \left(\delta_j * area_{ij}\right) + t_i + \epsilon_i\,, \qquad (1)$$

where i indexes suspect-stops or suspect-frisks (each suspect detained in a stop, and each separate time a suspect is detained will be a separate observation), $contraband_i$ is an indicator for whether contraband was discovered in the frisk, $after_i$ 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 ($year_{min}$ is 2016 for Chicago and 2015 for Philadelphia) and $area_{ij}$ is a dummy for the police region where the stop occurred (one of 70 "sectors" in Chicago and one of 66

"PSAs" in Philadelphia). t_i is a linear time trend. Suspect demographics are not included in our primary analysis as these are factors which police officers engaging in optimizing behavior may use to determine who to stop. The results from estimating a modified version of equation (1), which includes the age, race, and gender of the suspect, are reported in Table A2.

Table 2 reports the results from estimating equation (1) for pedestrian and vehicle frisks using the two separate events. As discussed in Section II, reasonable suspicion of a weapon is the legal justification for a frisk, but drugs constitute the bulk of what is recovered. Hence, we use two definitions of "hits": the discovery of a gun and the discovery of any type of contraband.

We focus our analysis initially on the protests (Table 2A). Across the two cities, only three of the results are statistically significant at conventional levels. However, all but one point in the expected direction - the vast decline in frisks corresponded to an increased hit rate consistent with optimizing behavior. The magnitude of the change is substantial relative to the low hit rates in each city. For example, the 2.6 percentage point increase post protests in the gun hit rate for pedestrians in Chicago is equivalent to 96% of the mean of 2.7% over the full period. In Philadelphia, the analogous 1.6 percentage point increase is equal to 89% of the mean.

The gun hit rate from vehicle searches in Philadelphia rose 1.5 percentage points post-protests (Table 2A). This is similar in magnitude but smaller relative to the mean hit rate of 3.1%. Meanwhile the increase in Chicago was only 0.4 percentage points. The results using all contraband as the main outcome are statistically significant at the 5% level with a 3.3 and 4.7 percentage point increase for vehicle stops in Chicago and Philadelphia. The only change which is negative is that for all contraband from pedestrian stops in Philadelphia. However, this is a relatively small drop and statistically insignificant.

The results using the pandemic onset 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 2.2 percentage points which is 129% of the mean for these dates. Meanwhile, in Philadelphia three of the results are negative although they are statistically insignificant at any reasonable level of significance. The only hit rate which moved in the expected direction was that for guns from vehicle stops which rose by a statistically significant 2.1 percentage points at the pandemic onset.

The pandemic findings are difficult to interpret for several reasons. One particular concern is that the frisk rate conditional on a stop actually increased during this period. Hence, while the threshold for stopping a suspect appears to have increased, the change in the threshold for search is ambiguous. This is because an officer can be expected to acquire additional information following a stop which informs their search decision. As a result, we cannot definitively conclude that under a model of optimal policing hit rates from search should have increased following this event.

To account for this, we examine contraband discovery as a share of all stops, rather than just frisks or searches. A large drop in stops should have the same implication as a large decline in frisks: 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 2016 to 2020 for Chicago and 2015 to 2020 for Philadelphia, for all stops using both the protests and pandemic events. The results for the protest period are similar to those in Table 2. Importantly, all point estimates for the pandemic period are now positive and a larger proportion are statistically significant.

The difference-in-difference approach in Tables 2 and 3 is our preferred specification because it accounts for seasonality. However, to ensure that prior year data are not driving the results, we run a single difference (before-after) comparison of hit rates around the protests, using only 2020 data. The results (Appendix Table A3) 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 increase in gun hit rates from vehicles stops in Philadelphia are now also statistically significant at the 5% level.

We now present results from an event study specification to analyze the timing of the impact (Figure 5):

$$contraband_i = \alpha + \sum_{t=-6}^{11}(\beta_{1t} * week_{it} * treat_i) + \sum_{t=-6}^{11}(\beta_{2t} * week_{it}) + \sum_{k=y_{min}}^{2020}(\gamma_k * year_{ik}) + \sum_{j=1}^{J}(\delta_j * area_{ij}) + t_i + \epsilon_i , \tag{2}$$

where variables are analogously defined as in equation 1. $week_{it}$ is a week dummy for the number of weeks which the stop occurred after 29 May of the given year for Chicago and 3 June for Philadelphia. We focus on the protest event for the event study as the decline in frisks was smaller and shorter lived following the pandemic onset (as seen in Figure 3). As this specification allows us to observe changes in the impact over time, we consider a longer window around the event than in our main specification presented in Table 1. What seems clear is that the hit rate of pedestrian frisks in Chicago and vehicle searches in both cities increased substantially after the protests and appears to have stayed at this higher level for the majority of the 3 months of post-protest data in Figure 5. The picture is less clear for pedestrian frisks in Philadelphia, which are noisier.

While these initial results are generally consistent with the implications of standard models of police stops, there may be a concern that other changes could have impacted the findings. 2020 was certainly not a typical year. Furthermore, the impact of the pandemic onset and the protests extended well beyond changes to police searches. Hence, it is possible that these results could be driven by other concurrent changes. The key threats to our identification include a change in the number or the composition of potential suspects, change in police deployment, or a change in police effort. We explore each of these possibilities below.

A. Policing Production Function by City, Stop, and Contraband Type

We return to the model of policing described in Section IIB and the policing production function introduced in Figure 2. 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 elasticity, which hasn't been previously explored. We do so here. In Figure 6 we present the empirical relationship between the number of frisks per day (per million residents) and the total number of successful frisks, as measured by pedestrian stops (Panel A) and vehicle stops (Panel B) respectively. In each panel, we further separate by Chicago and Philadelphia and by whether the "successful frisks" are defined as any contraband or gun discovery. [13]

For each of the 8 production function curves we have three observations: the origin, post-protests rates, and pre-protest rates. Thus, we may compute the hit rate elasticity with respect to frisk rate (noted by each curve). Given that the elasticities are computed from the minimum number of observations needed, these must be interpreted as broad averages. Cross-city comparisons, particularly for vehicle stops, must be interpreted with caution given that the higher search rates for the cities vary by almost a factor of 2. Elasticities closer to zero indicate frisks or

[13] We further separate the frisks by the race of the pedestrian or driver in Figure A1 in the Appendix. The qualitative conclusion remains robust.

searches closer to random (see the discussion in Section II. B), where there is no or low decline in hit rate as more frisks are performed. Given that frisks are costly, this suggests resources may be better allocated away from frisks, at least at the higher point on the curve.

For pedestrian stops, the lowest concavity is for contraband hit rates in Philadelphia, which appears as almost a straight line and has an elasticity of -0.0206. The low elasticity indicates that identification of individuals to frisk is little better than random. But how much better? One can quantify it by first constructing the counterfactual hits that would have occurred after the protest frisk decline if frisks or searches were completely random. In Appendix Table A4 we then report how many fewer hits these random searches would result in over a year, relative to the actual hit rate.

The second row of the table answers the question posed above: A completely random search process would result in only 5 fewer total contraband hits over an entire year in Philadelphia, less than a 1.5% decline. The elasticity in Philadelphia is over an order of magnitude smaller than the equivalent figure for Chicago, -1.075. This suggests that either costs of frisks are vastly different across the two cities or there may be gains in Philadelphia from a substantial reduction in frisks.

In Philadelphia, the gun hit rate elasticity at -0.6902 is over an order of magnitude greater than the general contraband hit rate, suggesting more useful information when focusing on this type of contraband. We see the same pattern in Chicago, where the gun hit rate elasticity at -1.4812 is almost 40% higher than that for overall contraband. In fact, this pattern is true for vehicle stops as well: given the city and stop type, there is always a higher elasticity for gun hit rate than all contraband. This is encouraging, as suspicion of weapons is the legal justification for a frisk or search, not suspicion of drugs or other contraband.

We now turn to the second panel of Figure 6, which reports analogous results for vehicle searches. As noted above, here the upper search rate in Philadelphia is about twice as high as for Chicago and thus it is harder to make cross-city comparisons. In Philadelphia, as with pedestrian frisks, vehicle searches for all contraband exhibit low concavity, with an elasticity of -0.2739. The gun hit rate is about twice as large at -0.4719. In Chicago, that rate is very similar, at -0.5053, but computed over a smaller range. That rate is higher than that for the all contraband hit rate, which is -0.4161. Still, with the exception of all contraband in Philadelphia, the elasticities are larger for pedestrian than vehicle stops. This suggests that vehicle stops are closer to random at least at the rate they are being performed in the cities.

Examining the policing production functions and computing the elasticities for the different combinations of city, stop type and contraband sheds further light on police stops. We find that when contraband is defined as firearms alone, elasticities are always higher than the analogous numbers when using all contraband. This shows the police respond more to information predicting firearm possession versus other types of contraband. This is encouraging, given the legal basis for frisks is suspicion of weapon possession. We also see that with one exception, elasticities are higher for pedestrian than the analogous vehicle stops. This suggests that officers have more predictive information from individuals on foot, rather than in vehicles. Finally, we see higher elasticities in Chicago than for the analogous measures in Philadelphia. This result should be interpreted with particular caution as the change in frisk rates differed in the two cities, and substantially for vehicle searches.

B. Testing for Changes in Police Deployment

As discussed in Section II.B, a key implication of policing models is that as an officer decreases the number of stops she makes, the hit rate of the marginal stop should increase. However, a decline in the total number of police stops in a city could result from a decline in the number of officers making stops rather than a decline in stops per officer. While the latter should unambiguously increase the hit rate of stops, the effect of the former depends on the spatial allocation of officers. For example, if an officer normally patrols a set area which overlaps with that of other officers, if those other officers are redeployed elsewhere during the protests, we would expect the remaining officer to encounter more high likelihood suspects and hence have a higher hit rate. If patrol territories do not overlap and suspects rarely move between patrol areas, then we would not expect an impact on hit rates from the redeployment of some officers.

While we do not have data on the patrol area of individual officers, we do observe identifiers of the individual officers making stops. Thus, we can determine whether the decline is on the *intensive* margin (stops per officer) or the *extensive* margin (number of officers making stops). Figure 7 reveals that following 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 decline is larger on the extensive margin than on the intensive margin. In Philadelphia, the number of officer pairs making stops fell by about 75% while the decline in stops per officer pair was about 35%.[14] The change in Philadelphia following the pandemic onset is particularly interesting, with no decline on the intensive margin and the entire fall in stops occurring through the extensive margin.

This difference in the changes to stops per officer pair could explain the differing results in Table 2, which showed that in Philadelphia hit rates from searches generally slightly declined in the period immediately after the pandemic onset but rose after the protests. As discussed above, changes in the extensive margin have an ambiguous impact on hit rates, depending on deployment of officers and the distribution of suspects. Meanwhile, declines on the intensive margin should unequivocally increase hit rates, which is what we observe following the protests.

One way to focus on the impact of the change in the intensive margin is by restricting analysis to the subset of officers who made stops in the 6 weeks after 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 for consistency with policing models; as individual officers become more selective in stops and frisks, their hit rates should increase. In addition, we estimate equation (1) including *officer fixed effects* with the results reported in Table A5. Again, the estimated change in hit rates remains positive and of a similar magnitude although the increased standards errors means that very few of the estimates are statistically significant.

We also test whether officers with higher hit rates in the pre-protest period were more likely to be deployed to make stops during the post-protest period. Specifically, we rank the officers with at least ten stops in the pre-protest period according to their hit rates and then regress an indicator variable for whether the officer was active in the post-protest period on their pre-protest quartile ranking of hit rates. We find that the coefficient estimate for the ranking is not statistically significant. This suggests that police Departments did not select officers in the post-protest period based on their skill in detecting contraband. We also rule out a large change in the

[14] In both cities, the majority of stops in our data involve two officers patrolling together. We expect that these officers coordinate their decisions about whether to stop and frisk people, so we conduct our analysis at the officer pair level. For stops involving a single officer, we treat that officer as equivalent to a "pair".

composition of employed officers following these events. In Figure A2, we plot the monthly number of officer resignations from the Chicago Police Department. Neither event caused a meaningful increase in the number of resignations relative to the size of the Chicago Police Department, which employs roughly 12,000 sworn officers.

Another concern related to officer deployment is that changes in the geographic distribution of stops may have driven the increased hit rates. If a larger proportion of stops occur in areas of the city where contraband carry rates are higher, then hit rates could increase even without individual officers engaging in optimizing behavior. Such an effect would instead potentially indicate optimizing behavior on the part of police management. To control for geographic changes, we repeat our analysis constructing the "After*Treat" data by sampling with replacement from the raw data and requiring that the distribution of stops across police districts matches that observed in the "Before" period of 2020. We use the data for the prior years and the "Before" period of 2020 without any modification. The model in equation (1) is then estimated and the results are reported in Table A6. The results are once again qualitatively the same, with hit rates increasing after the protests in both cities and after the pandemic onset in Chicago.

One concern about the sharp decline in stops per officer pair in Figure 7 is that this may indicate a reduction in police effort. A decrease in effort could cause a decline in the hit rate *ceteris paribus* as it may cause officers to move from optimizing behavior towards something approaching random search. This would imply that the diminishing returns to additional stops are greater than those we estimate. It is not possible to completely rule out the possibility that effort fell as a result of the protests. One additional piece of information that casts doubt on the possibility of reduced officer effort is provided in Table 5. This table reports results from estimating equation 1 in Philadelphia, but where the dependent variable is a dummy for whether the stop or frisk lacked legal foundation. An illegal stop would suggest that the officer has not correctly considered whether they have legal foundation for the stop and thus we expect that the proportion of illegal stops would be negatively correlated with police officer effort. In both cases, the point estimates show a decline in the proportion of illegal stops and frisks. Thus, we believe this is evidence against the hypothesis that officers were expending less effort in determining who to stop/frisk. Hence, we believe it is unlikely that a large change in police effort has substantially impacted the results. Additionally, to the extent that there was a reduction in police effort, this would bias our results downwards and would imply an even greater increase in hit rates from a reduction in stops.

C. Testing for Changes in Suspect Population

The predicted effect on hit rates of a greater number of individuals on the street depends upon the model of policing. Greater foot traffic increases the number of potential suspects, so if observable characteristics are informative then officers would have a larger pool of higher likelihood suspects and should thus have a greater hit rate. However, if the additional individuals carry contraband at a lower rate this would lower the proportion of individuals with contraband. Hence, under a random model of policing or one where officers only have a weak ability to predict contraband carriers based on observables, an increase in potential suspects could cause a decrease in the hit rate.

Figure 8 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. At the onset of the pandemic there is a sudden, deep fall in mobility and then a slow recovery beginning in April 2020. The drastic change in foot traffic at the pandemic onset is a clear confounding factor when analyzing the impact of the decline in police stops during this period. As discussed above, the bias introduced by this change in mobility could be either upward or downward, depending on the model of policing and assumptions about the potential change in the composition of people on the street following the pandemic onset. This does not apply to the period around the protests when there was a smooth but relatively modest increase in mobility measures.

We analyze the age distribution of motor vehicle crash patients as a proxy to investigate changes in the demographics of those travelling by car. This data is recorded by the CPD in Chicago and in the PTOS trauma database in Philadelphia, and plotted in Figure 9. The top panel for each city shows changes around the pandemic onset and the bottom shows changes around the protest period. In each figure, the blue dashed bars measure the density prior to the event and the red solid bars afterward.

There are substantial changes around the pandemic onset. The differences are particularly large in Philadelphia, with a substantial tightening in the age distribution. The change following the 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 for changes in the age distribution around the protests (Appendix Table A7) fails to reject the null of no change in both cities.

Another potential cause for changes in the suspect population over this period is jail releases. These releases were a common policy to reduce crowding in jails at the beginning of the COVID-19 pandemic. With respect to the protest period, it is very unlikely that this would affect our results given that these releases predominantly occurred between late March and early May. Jail populations were quite stable in Philadelphia and actually rose slightly in Chicago over the relevant time period following the protests. The timing of these releases does overlap with the period for our analysis related to the pandemic onset. While it cannot be ruled out that this impacted the pool of potential suspects, it should be noted that only ~1000 residents were released during this period in each city. Thus, we do not expect this to have had a meaningful impact on the underlying rate of criminality in the population.

Overall, while it is doubtful that jail releases had an effect on foot or vehicular traffic, it seems probable that the reduced mobility during the initial period of the pandemic changed the potential suspect pool markedly. For this reason, we focus on changes following the protests for the rest of the paper. While there were certainly localized changes in potential suspects as protests or looting occurred, overall mobility patterns changed smoothly and at least one measure of vehicular travel did not change appreciably. We further explore potential changes in the contraband carry rate around the time of the protests below.

D. Change in Crime

While we have examined potential changes in the composition of pedestrians and drivers in Section IV.A above, it is still possible that there could be a change in the contraband carry rate unrelated to changes in overall numbers or demographics.[15] If a larger proportion of individuals started engaging in criminal activity, the proportion of individuals on the street carrying contraband would likely increase. This would cause hit rates to rise even in a setting of random policing as hit rates should equal the contraband carry rate under such a model.

---

[15] See Rivera and Ba (2023) for a detailed discussion of the risk of changes in civilian behaviour when using scandals or protests to study the effect of changes in police behaviour.

To investigate this, we estimate equation 1 using contemporaneous crime reports as the dependent variable (Table 6). We use crime reports as a measure of the number of crimes committed. While it is possible that reporting behavior changed in this period, using data from the National Crime Victimization Survey, Ang et al. 2023 find a decline of just 1.5 reports per 100 offenses.[16] Although changes vary substantially across crime types, most policing tends to target 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. Thus, it appears highly unlikely that changes in criminal propensity among the suspect population is responsible for the rise in hit rates which we observe in Philadelphia. If anything, criminality may have fallen slightly in this period, without which the hit rate increase may have been even larger than what we observed.

However, Chicago experienced a large and statistically significant increase in violent crime in the period after the protests. Overall violent crime increased by 20% and shootings increased by over 45%. Given that an increase in shootings would likely be associated with an increase in the gun carrying rate, this change in civilian behavior could have caused an increase in gun hit rates.

To identify whether a change in carry rates drove the increase in hit rates in Chicago, we estimate the change in hit rates for a subset of Chicago where shootings did not increase substantially. The eight police districts we look at are shown in black in Figure 10 and consist of an almost contiguous subset of the city which represents approximately half of the police stops observed in our data. 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 increases in overall violent crime and shootings specifically are much smaller in magnitude (only 11.5% and 6% respectively) and statistically insignificant. Despite this, we still observe a substantial increase in the hit rate for guns from pedestrian frisks, which increased by 4 percentage points across these districts. However, larger standard errors due to the reduced number of observations mean that this increase is no longer statistically significant. Thus, we conclude that the increase in hit rates

[16] They find a substantially larger decrease in the ratio of 911 calls to Shotspotter detections. However, this decrease may be at the intensive margin rather than extensive margin i.e., a reduction in calls per shooting rather than a reduction in the number of shootings that police are notified of. In addition, Shotspotter has been criticised for issues such as struggling to separate fireworks from gunfire (Carr & Doleac, 2016), which is particularly relevant given the wave of fireworks during this period. Due to accuracy concerns, its use in Chicago is being discontinued.

which we observe cannot be explained solely by a change in the underlying rate of criminality in the population following the protests.

E. Race Effects

Our analysis so far has focused on overall changes in hit rates. But hit rate tests are most frequently employed to assess racial disparities in policing. Hence, it is worthwhile separating out the changes we observe by the race of the suspect. In the discussion of Table 1, we noted the fairly substantial decline in the Black share of stops in Philadelphia following the protests, both pedestrian and vehicle. The fact that the subject of the protests was racially disparate policing could have been a reason for this change. Thus, under an optimizing model of policing, it seems reasonable to interpret this change as an increase in the threshold for stopping or frisking a Black pedestrian or motorist relative to White pedestrians. Hence, in Philadelphia we should expect an increase in the hit rate for Black individuals relative to Whites. Table 8 reports results from the triple difference specification in equation 3

$$contraband_i = \alpha + \beta_1 after_i * treat_i * Black_i + \beta_2 after_i * treat_i + \beta_3 after_i * Black_i + \beta_4 treat_i * Black_i + \beta_5 Black_i + \beta_6 after_i + \sum_{k=y_{min}}^{2020} (\gamma_k * year_{ik}) + \sum_{j=1}^{J} (\delta_j * area_{ij}) + t_i + \epsilon_i \quad (3)$$

The results here do not provide clear support for this hypothesis. Even as the share of Black motorists stopped in Philadelphia declines; the relative hit rate also drops appreciably for vehicle stops. Meanwhile, among pedestrian stops the change is quite small for the gun hit rate, and positive and large for all contraband, as would be expected. There are many potential reasons. One possibility is that the underlying contraband carrying rate among Black drivers and passengers may have declined relative to Whites during this period of uncertainty around the protests. Second, concerns of racial bias may have caused police officers to become more likely to issue warnings to Black individuals rather than formally recording contrabands. This would serve to lower the Black hit rate observed in our data. However, neither explanation would justify the difference between the changes for pedestrian hit rates relative to vehicle hit rates. It is also possible that officers misreport race for a proportion of failed stops and that the incentives for misreporting change in response to the protests. To explain the observed change in vehicle hit rates, we would require a large proportion of failed stops of Black motorists being misclassified as involving a White driver and that the costs for misreporting Black motorists as White increased as a result of the protests. However, Luh finds that in her context of Texas patrol stops, the majority of misclassification occurred with respect to Hispanic motorists. Further, this explanation would also be unable to explain the differences in the observed changes for pedestrian stops relative to vehicle stops.

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 limits the available data. A longer window increases the number of observations but weakens the focus. In Table A8, we present results with varying time windows to assess the sensitivity of our results to precise timing.

Two other timing issues are examined, besides the length of the event 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. This is done to ensure that the results are not driven by high hit rates during the protest activities. Here we once again find the results to be consistent with our main findings.

B. Discussion

While most of the main specifications show an increased hit rate as stops and frisks dropped tremendously around the protests, the hit rate declines slightly for pedestrian stops in

Philadelphia in the “all contraband” specification, driven by drug discoveries. At the same time, the hit rate rises substantially for all contraband from vehicle stops. These results may indicate that marginal frisks are predominantly aimed at drug recovery while higher priority frisks are aimed at detecting weapons. This would align with the legal justification of frisks being reasonable suspicion that the individual is armed..

As mentioned above, a legal pedestrian stop requires that a police officer has reasonable suspicion that the person to be stopped is currently or about to be engaged in criminal activity. To then frisk the individual, the officer must have reasonable suspicion that they are armed. Notably, 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 exercised more care to ensure the legality of their actions. This is supported by Table 5, where the proportion of illegal frisks decreased. Hence, we would expect that police reduced the number of frisks conducted solely to discover drugs (which are illegal frisks). 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 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. Thus, the drop in the overall search rate still dominates, resulting in a higher hit rate.

Perhaps the greatest concern is that gun carry rates may have increased sharply after the protests, and this is responsible for the increased hit rates. Section IV.C presents the results of several additional tests to rule out this possibility. In Philadelphia, we find that most crimes drop after the protests (Table 6), including shootings, which one would expect to rise with gun carry rates. In Chicago, where crime rates rose overall, we show that the areas of the city with the lowest growth in the rate of shootings still have an increased hit rate (Table 7B). In the absence of a direct measure of gun carry rates, one can't entirely rule out a sharp change, but all the evidence assembled indicates this is unlikely to be the sole reason for the abrupt growth in hit rates.

It is worth considering the apparent contradiction between our findings and those of Feigenberg and Miller (2022). As mentioned in Section II, Feigenberg and Miller (2022) find no relationship between search rates and contraband hit rates in Texas highway patrol stops. Our analysis differs from theirs in the source of variation and in the nature of the stops. They use between-officer variation to proxy for within-officer changes. However, as they note, this requires assuming that the screening ability of troopers is independent of their propensity to search motorists. They show suggestive evidence in support of this assumption. However, if it does not in fact hold in their setting, this could drive the difference in their results from *between-officer* variation and ours from *within-officer* variation.

There are substantial differences in the nature of stops between our settings. In making a highway stop of cars sometimes zooming along at 90 miles per hour, troopers have very little information to make their decision. Meanwhile, in our setting of inner-city stops, officers have more time to observe the behavior and characteristics of an individual or vehicle before deciding whether to engage in a stop. This is particularly true for pedestrian stops, which, as highlighted in Figure 6, is where we observe the greatest diminishing returns. Indeed, we show in Section IV. A that even within our setting, hit rate elasticities are mostly higher for pedestrian stops and, in general, closer to zero for vehicles. Greater information should allow officers to more accurately estimate the probability an individual is carrying contraband and thus engage in behavior consistent with optimizing models. This suggests that the models of police stops need to be sensitive to the context they describe.

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 against police brutality following George Floyd's killing. 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 (2001) and Anwar and Fang (2006) 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 (2006).

We are also able to estimate elasticities for the policing production function. We find that when contraband is defined as firearms alone, elasticities are always higher than when defined as all contraband (including drugs), which suggest that the police respond more to information predicting firearm possession versus other types of contraband. We also find that with one exception, elasticities are higher for pedestrian than the analogous vehicle stops, which suggests that officers have more predictive information from individuals on foot, rather than in vehicles.

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. (2014) The Law and Economics of Stop-and-Frisk, 46 *Loyola Law Review* 369

Abrams, D. S., Fang, H., & Goonetilleke, P. (2022). Police Frisks. *AEA Papers and Proceedings* 112, 178-183.

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

Ang, D., Bencsik, P., Bruhn, J., & Derenoncourt, E. (2023). Shots fired: Crime and community engagement with law enforcement after high-profile acts of police violence. *Benjamin H. Griswold III, Class of 1933, Center for Economic Policy Studies, Department of Economics, Princeton University.*

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

Arnold, D., Dobbie, W. & Yang, C. S. (2018), 'Racial Bias in Bail Decisions', Quarterly Journal of Economics 133(4), 1885–1932.

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

Bacher-Hicks, A., & de la Campa, E. (2020). The impact of New York City's Stop and Frisk program on crime: The case of police commanders.

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

Bohren, J. A., K. Haggag, A. Imas, and D. G. Pope (2020): "Inaccurate Statistical Discrimination: An Identification Problem," NBER Working Paper No. 25935.

Carr, J., & Doleac, J. L. (2016). The geography, incidence, and underreporting of gun violence: new evidence using ShotSpotter data. *Incidence, and Underreporting of Gun Violence: New Evidence Using Shotspotter Data (April 26, 2016)*.

Chalfin, A. and McCrary, J. (2017) "Criminal Deterrence: A Review of the Literature" *Journal of Economic Literature*, 55(1).

City of Chicago. 2020. "Crimes - 2001 to Present". Chicago Data Portal. https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-Present/ijzp-q8t2 (accessed January 4, 2021).

Chicago Police Department. 2021. "ISR Data". https://home.chicagopolice.org/statistics-data/isr-data/ (accessed January 4, 2021).

Chicago Police Department. 2020. "Chicago Shootings-2015-2020".

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. (2022) Would Eliminating Racial Disparities in Motor Vehicle Searches have Efficiency Costs?, *The Quarterly Journal of Economics*, Volume 137, Issue 1, February 2022, Pages 49–113, https://doi.org/10.1093/qje/qjab018

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

Gelman, A., Fagan, J., & Kiss, A. (2007). An analysis of the New York City Police Department's "stop-and-frisk" policy in the context of claims of racial bias. *Journal of the American Statistical Association*, 102, 813-823.10.1198/016214506000001040

Gonçalves, F. M., & Mello, S. (2023). Police discretion and public safety (No. w31678). *National Bureau of Economic Research.* https://doi.org/10.3386/w31678

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

Hull, P. (2021), What marginal outcome tests can tell us about racially biased decision-making, Working paper. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id3795454

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

Luh, E. (2022). Not so black and white: uncovering racial bias through systematically misreported trooper reports. *Working Paper, University of Michigan Criminal Justice Administrative Records System.*

MacDonald, J. & Anthony A. Braga (2019) Did Post-Floyd et al. Reforms Reduce Racial Disparities in NYPD Stop, Question, and Frisk Practices? An Exploratory Analysis Using External and Internal Benchmarks, *Justice Quarterly*, 36:5, 954-983, DOI: 10.1080/07418825.2018.1427278

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

City of Philadelphia. 2020. "incidents_part1_part2". Open Data Philly.

https://opendataphilly.org/datasets/crime-incidents/ (accessed January 7, 2021).

Philadelphia Police Department. 2021. “Vehicle and Pedestrian Investigations".

City of Philadelphia. 2020. “shootings". Open Data Philly. https://opendataphilly.org/datasets/shooting-victims/ (accessed December 16, 2021).

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

Rivera, R. G., & Ba, B. A. (2023). The Effect of Police Oversight on Crime and Misconduct Allegations: Evidence from Chicago. *Review of Economics and Statistics*, 79(1), 1-45. https://doi.org/10.1162/rest_a_01377

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 1. An Illustration of the Optimizing Model of Police Behavior

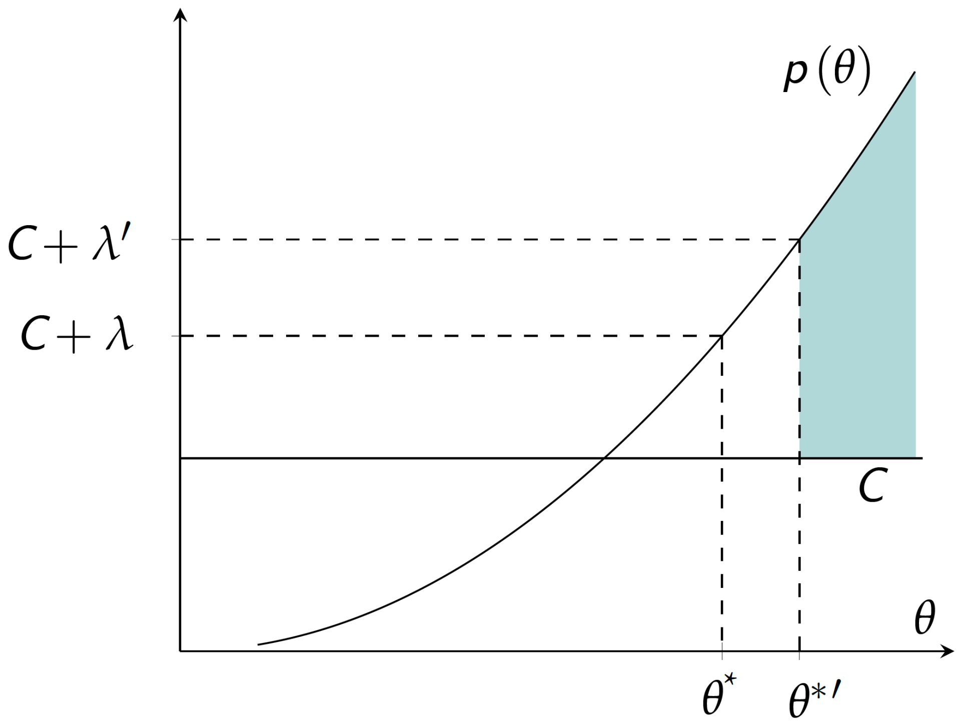
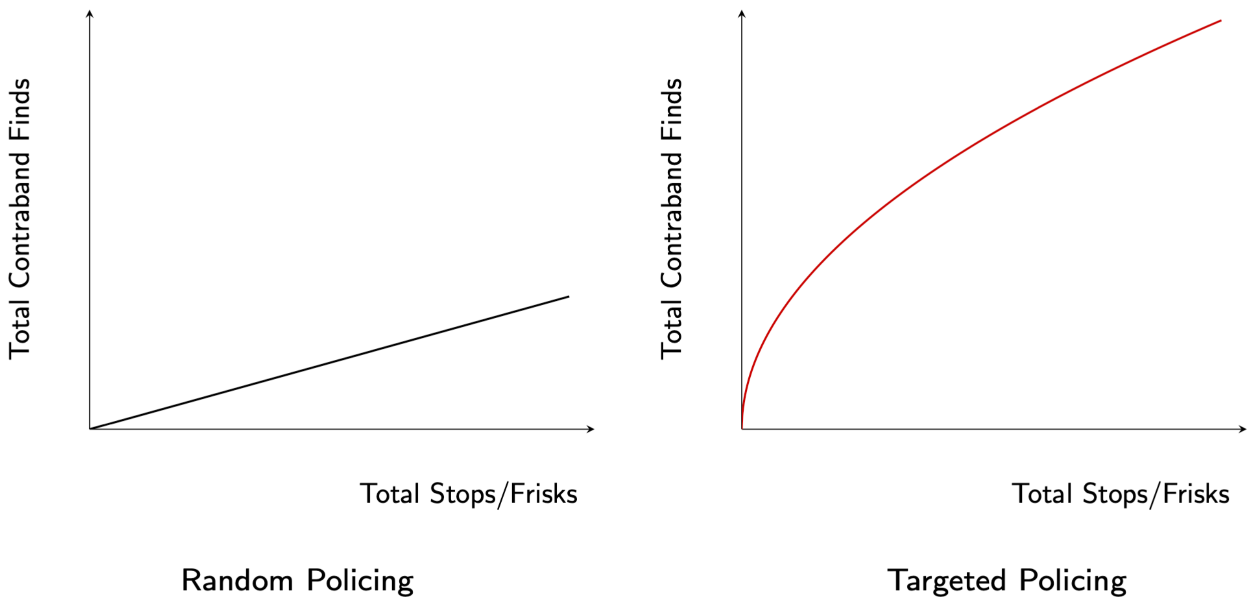
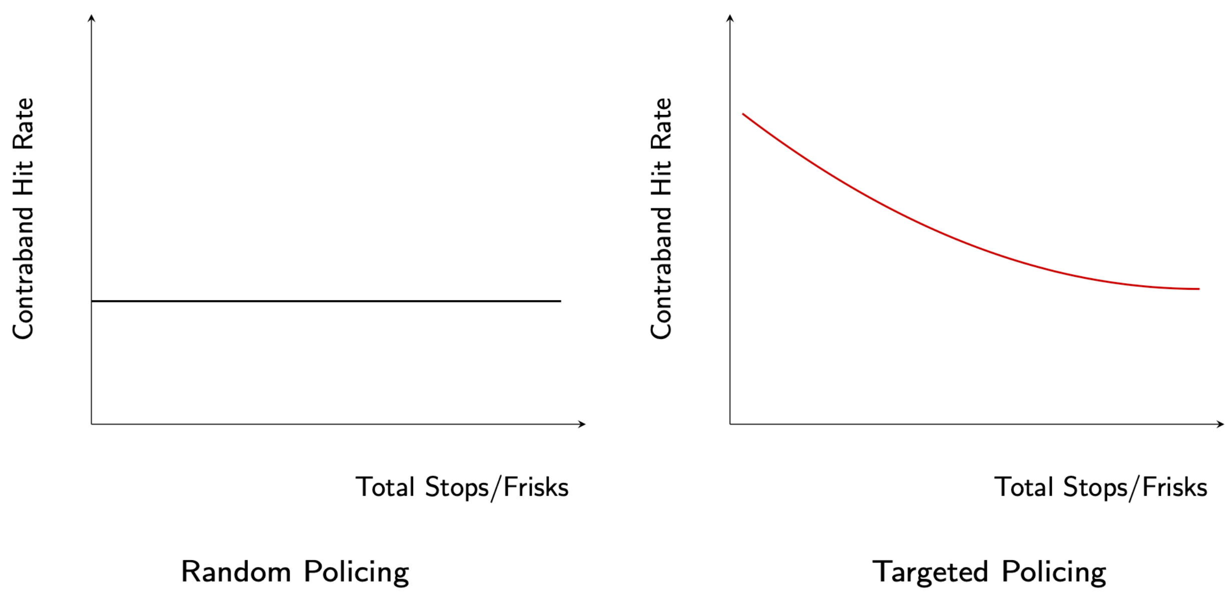


Figure 2: Policing Production Function for Random vs. Targeted Policing



Panel A: Total Stops/Frisks and Total Contraband Finds



Panel B: Total Stops/Frisks and Contraband Hit Rate

Figure 3a

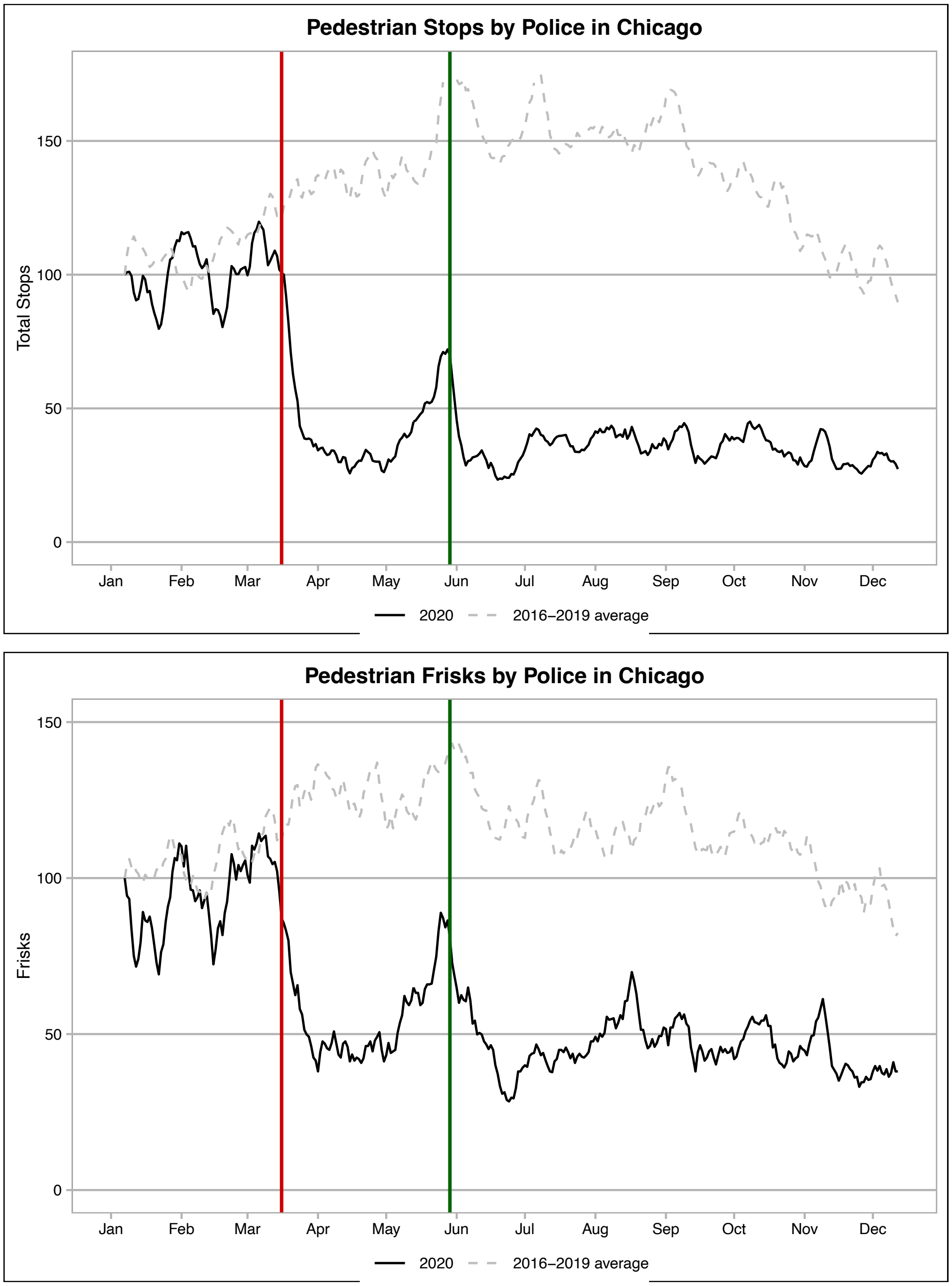
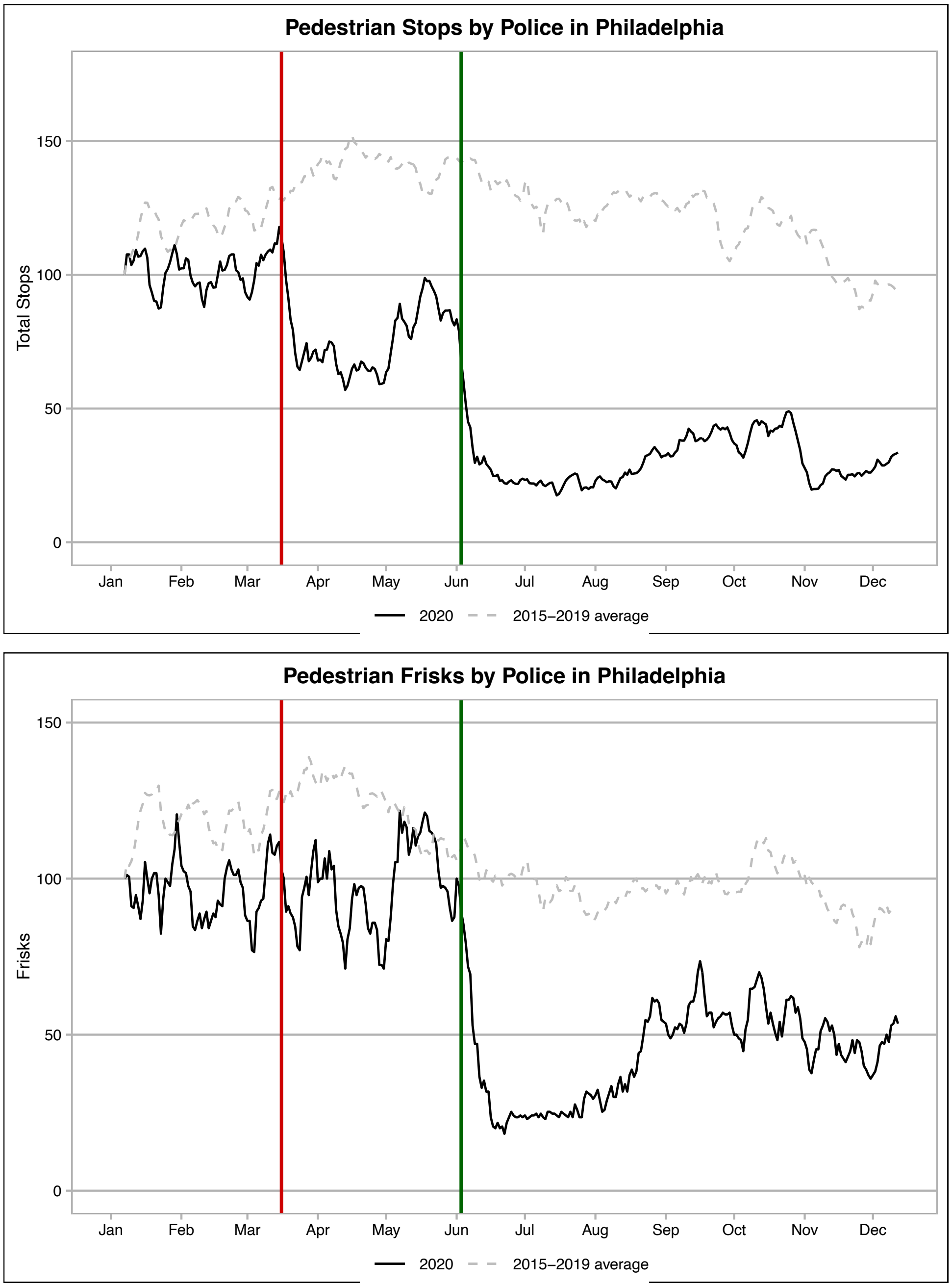
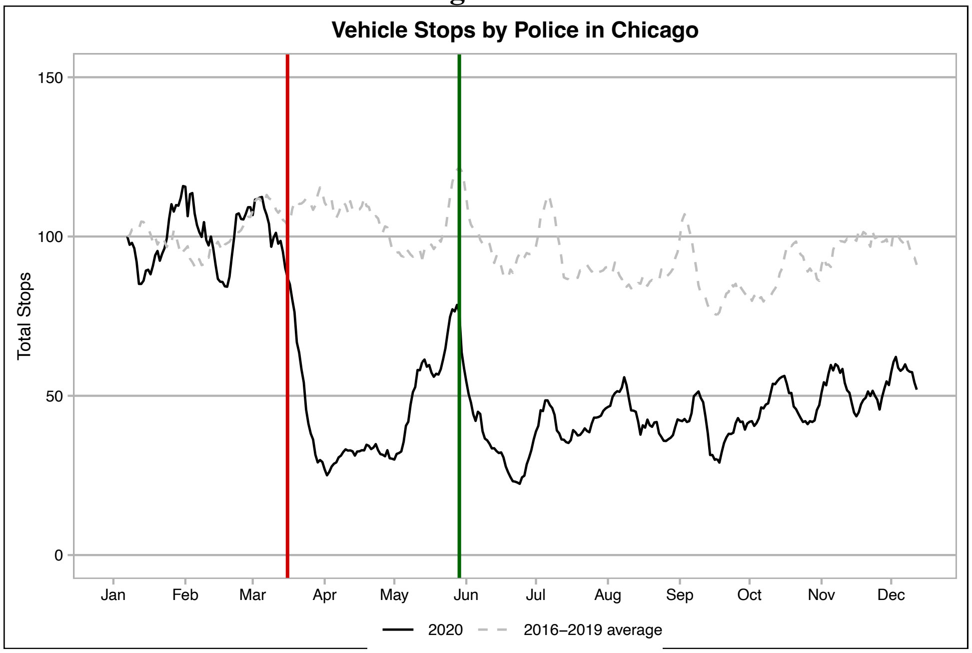

Pedestrian Stops by Police in Chicago
Total Stops
0
50
100
150
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
2020
2016–2019 average
Pedestrian Frisks by Police in Chicago
Frisks
0
50
100
150
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
2020
2016–2019 average

Figure 3b



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 value of which is normalized to 100. 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 4a


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
2016–2019 average

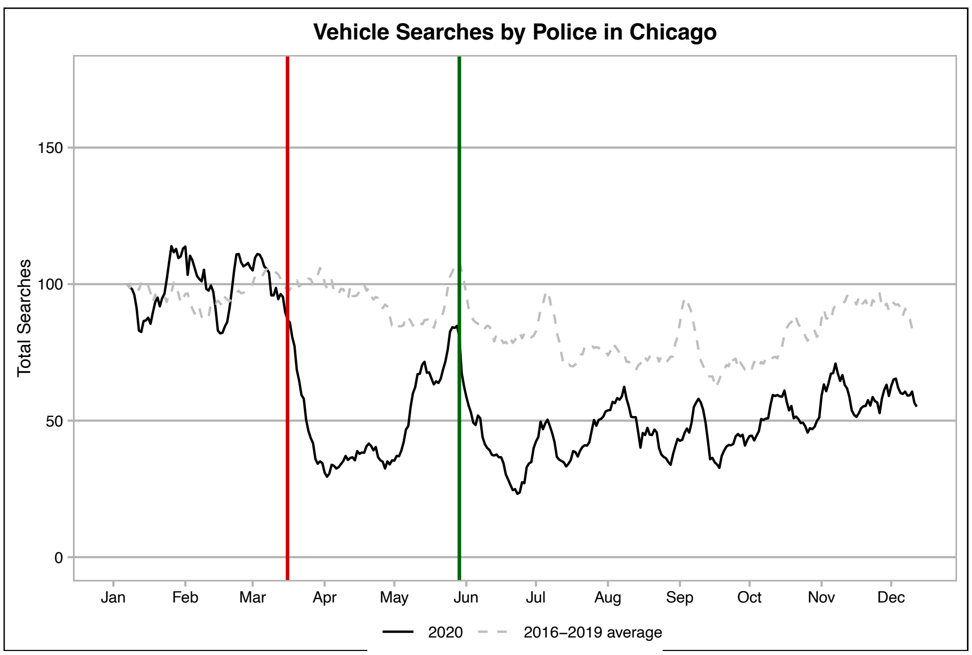
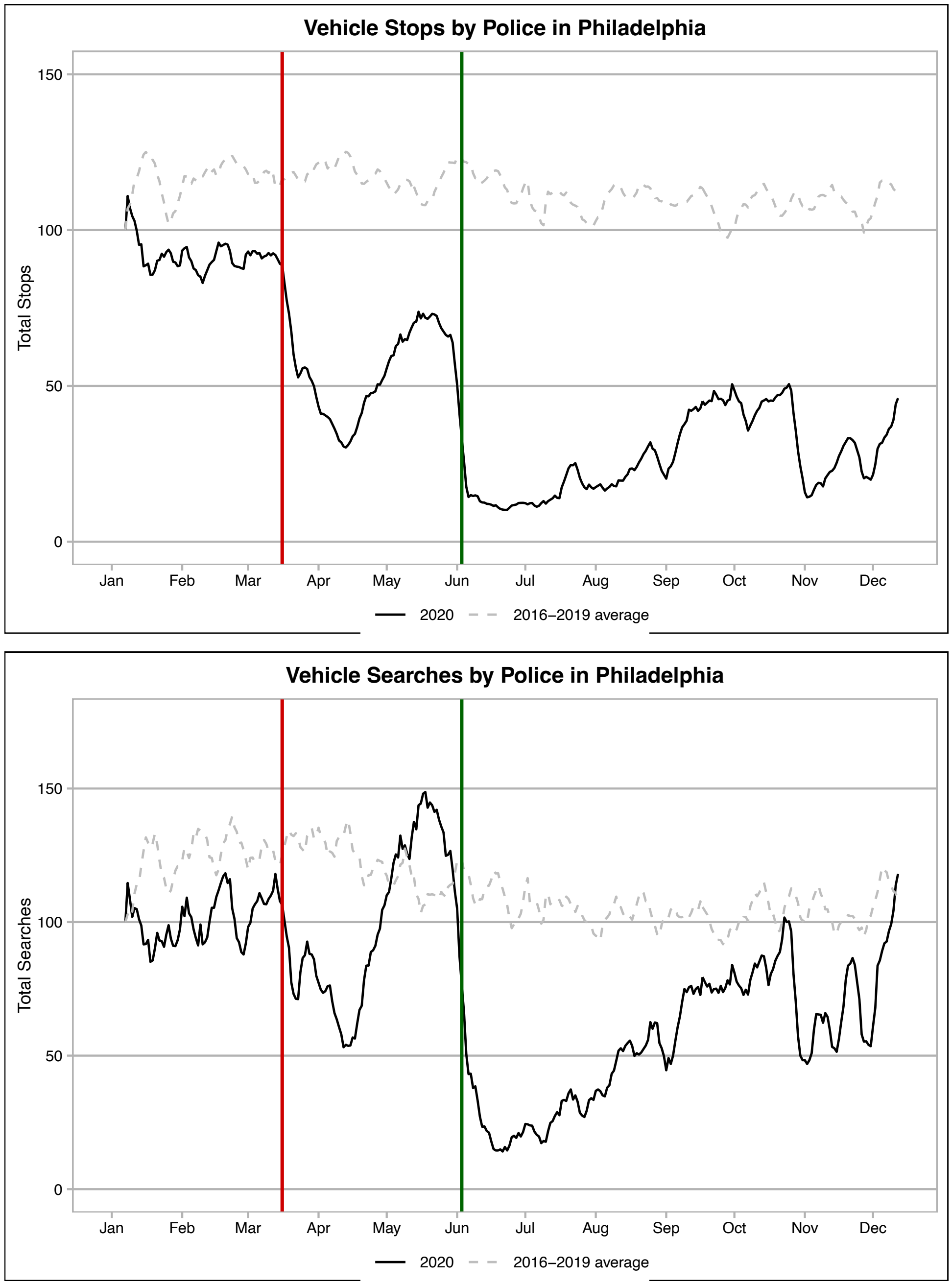
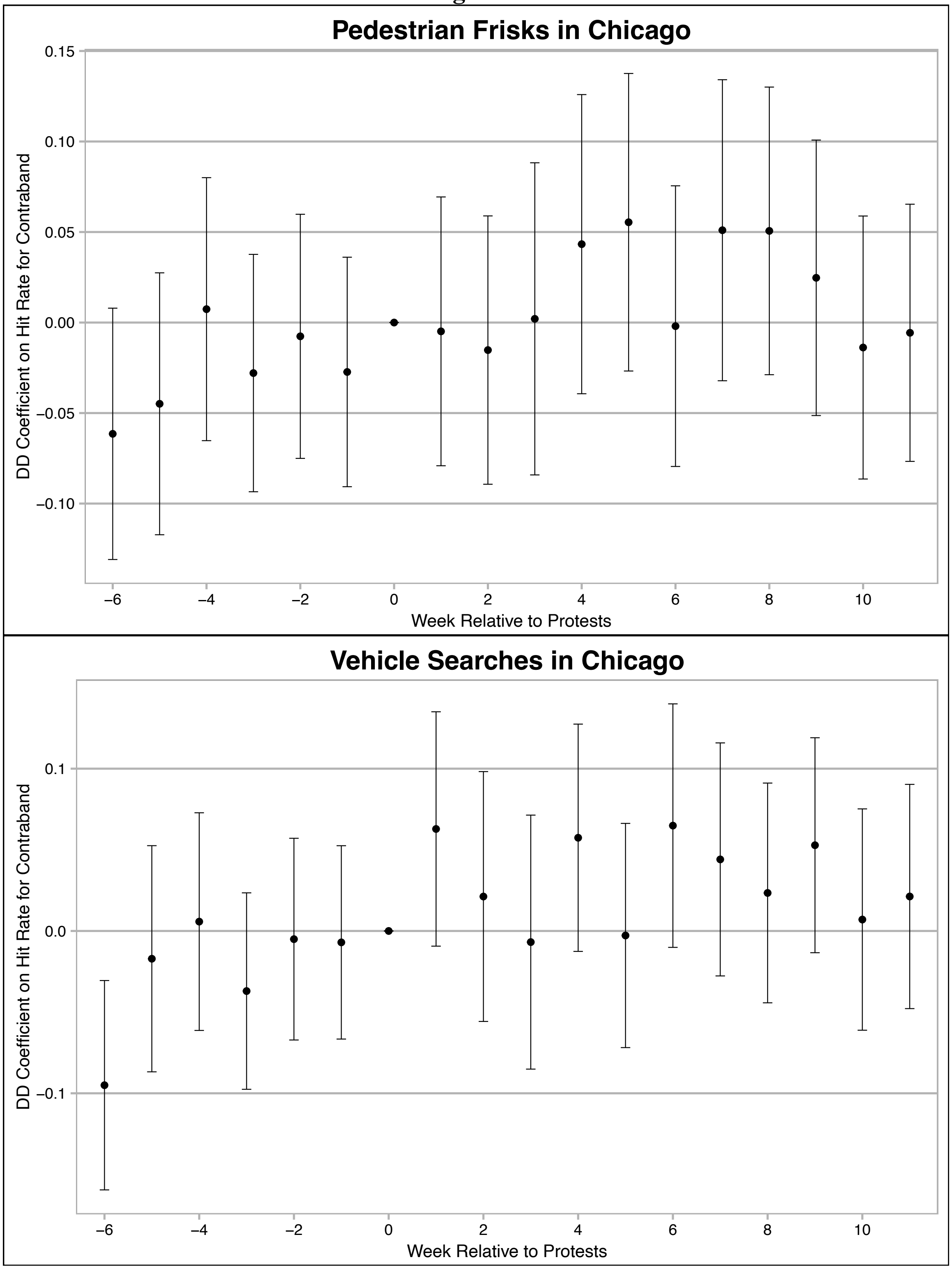

Vehicle Searches by Police in Chicago
150
100
50
0
Total Searches
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
2020
2016–2019 average

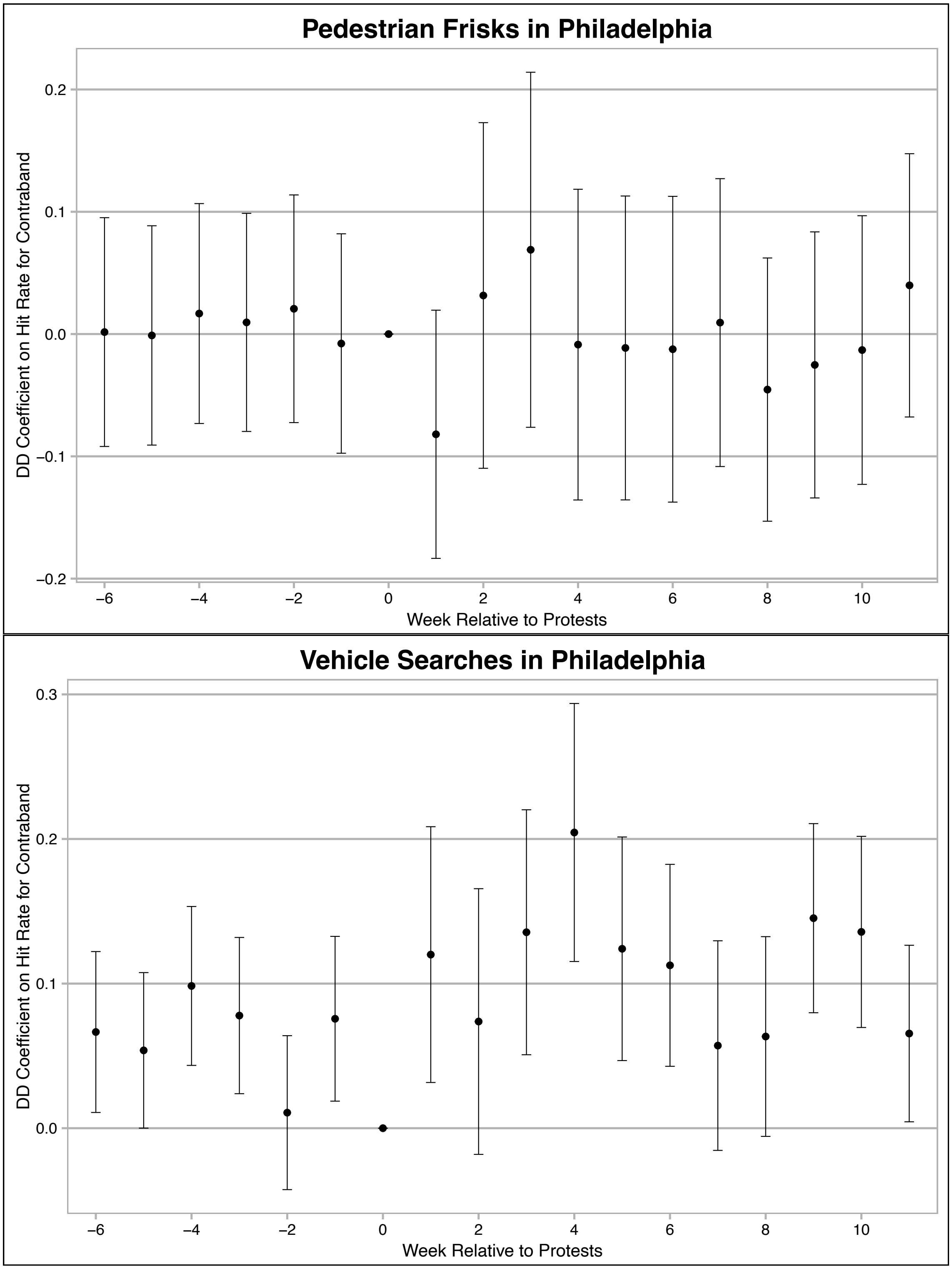
Figure 4b



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 value of which is normalized to 100. 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 5


Pedestrian Frisks in Chicago
DD Coefficient on Hit Rate for Contraband
0.15
0.10
0.05
0.00
−0.05
−0.10
−6
−4
−2
0
2
4
6
8
10
Week Relative to Protests
Vehicle Searches in Chicago
DD Coefficient on Hit Rate for Contraband
0.1
0.0
−0.1
−6
−4
−2
0
2
4
6
8
10
Week Relative to Protests



Event study: Hit Rate for finding contraband conditional on frisk/search in 2020 relative to 2016-2019 for Chicago 2015-2019 for Philadelphia. 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 2016-2019 for Chicago and 2016-2019. Year and police region fixed effects as well as a time trend are included. The vertical lines for each coefficient show the 95% confidence intervals from robust standard errors clustered at the region level

Figure 6: Relationship between Number of Frisks and Successful Frisks

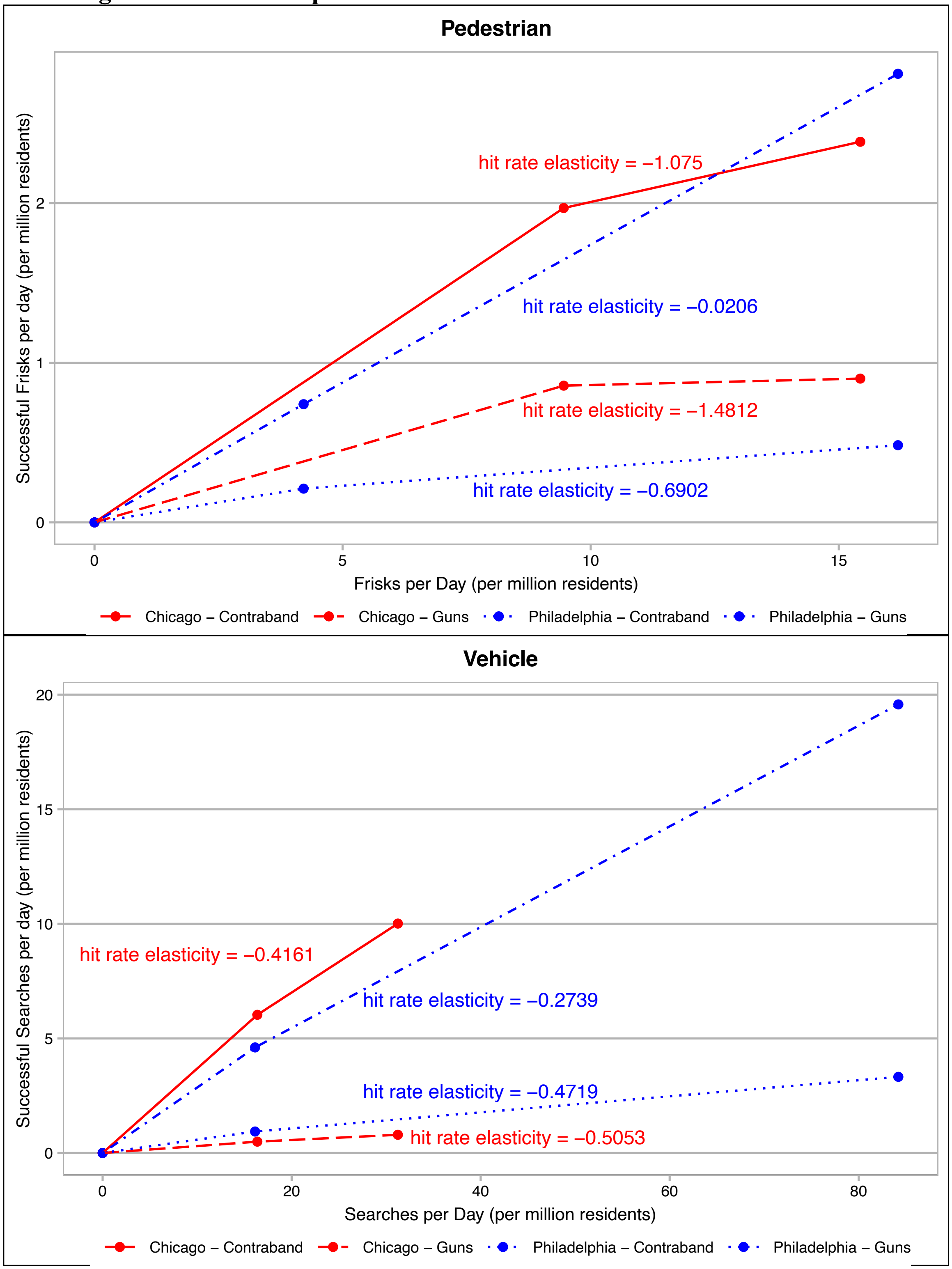
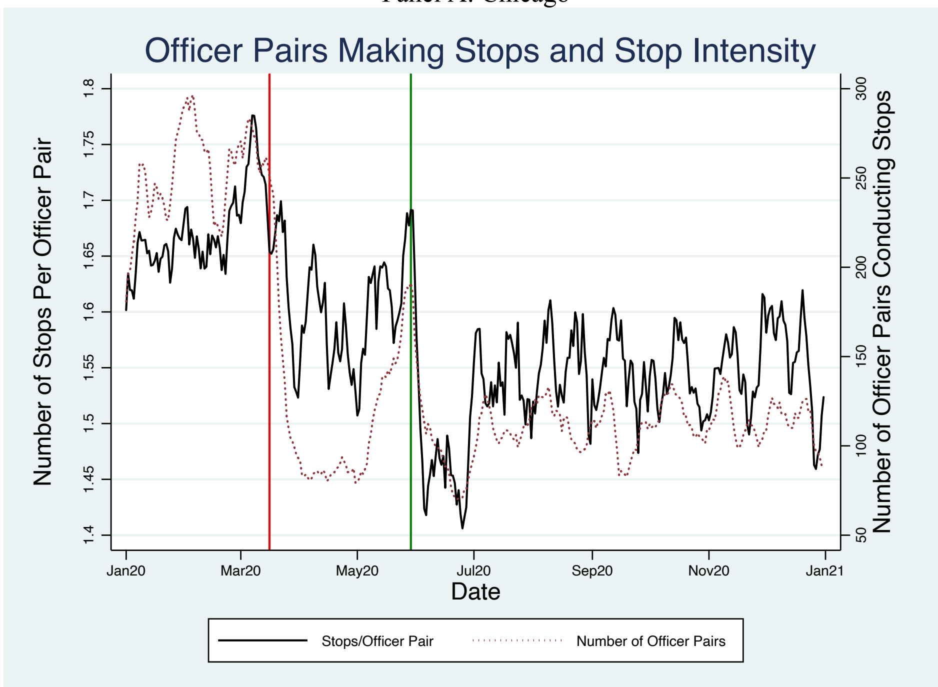


Figure 7

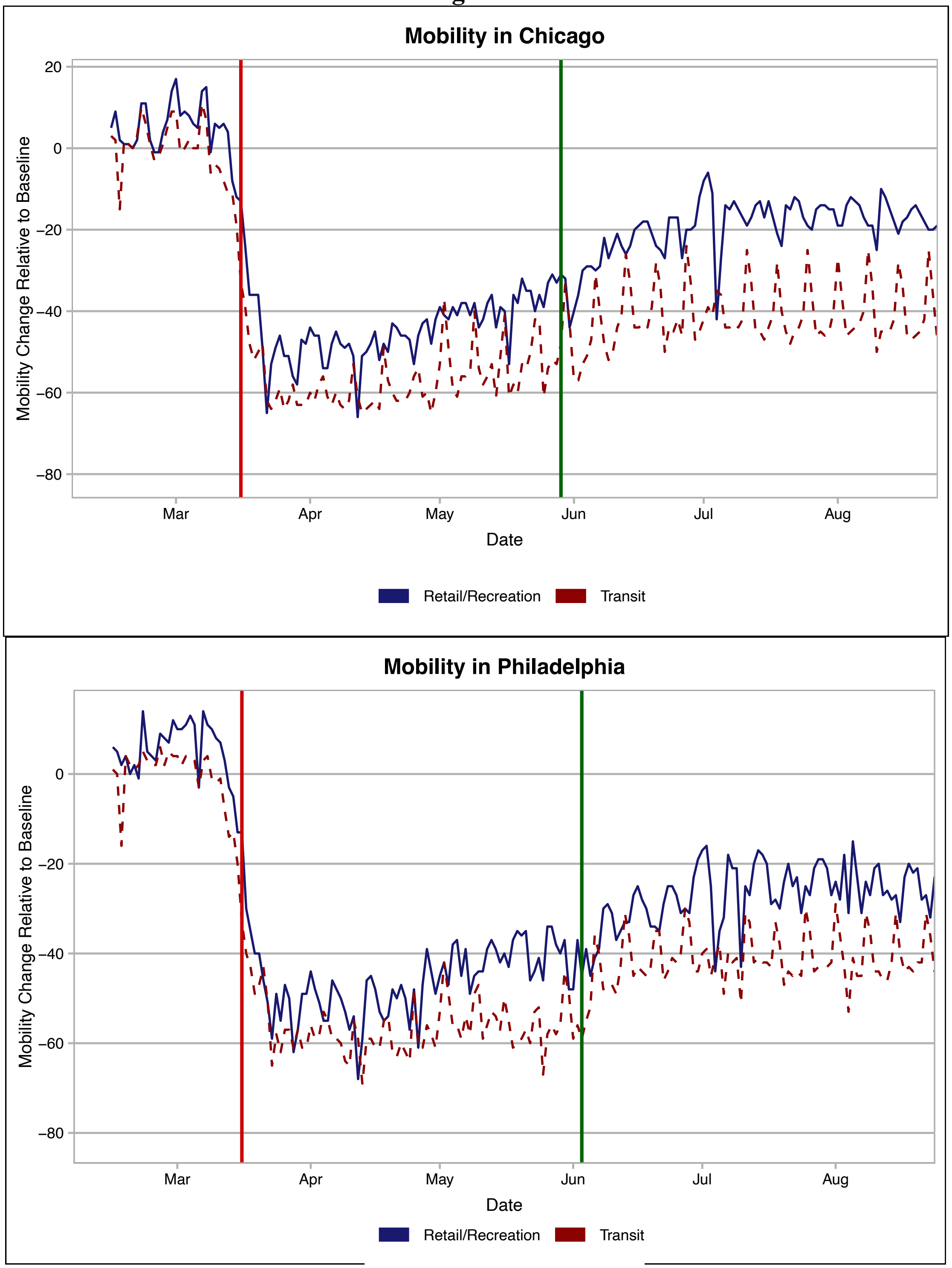
Panel A: Chicago



Panel B: Philadelphia

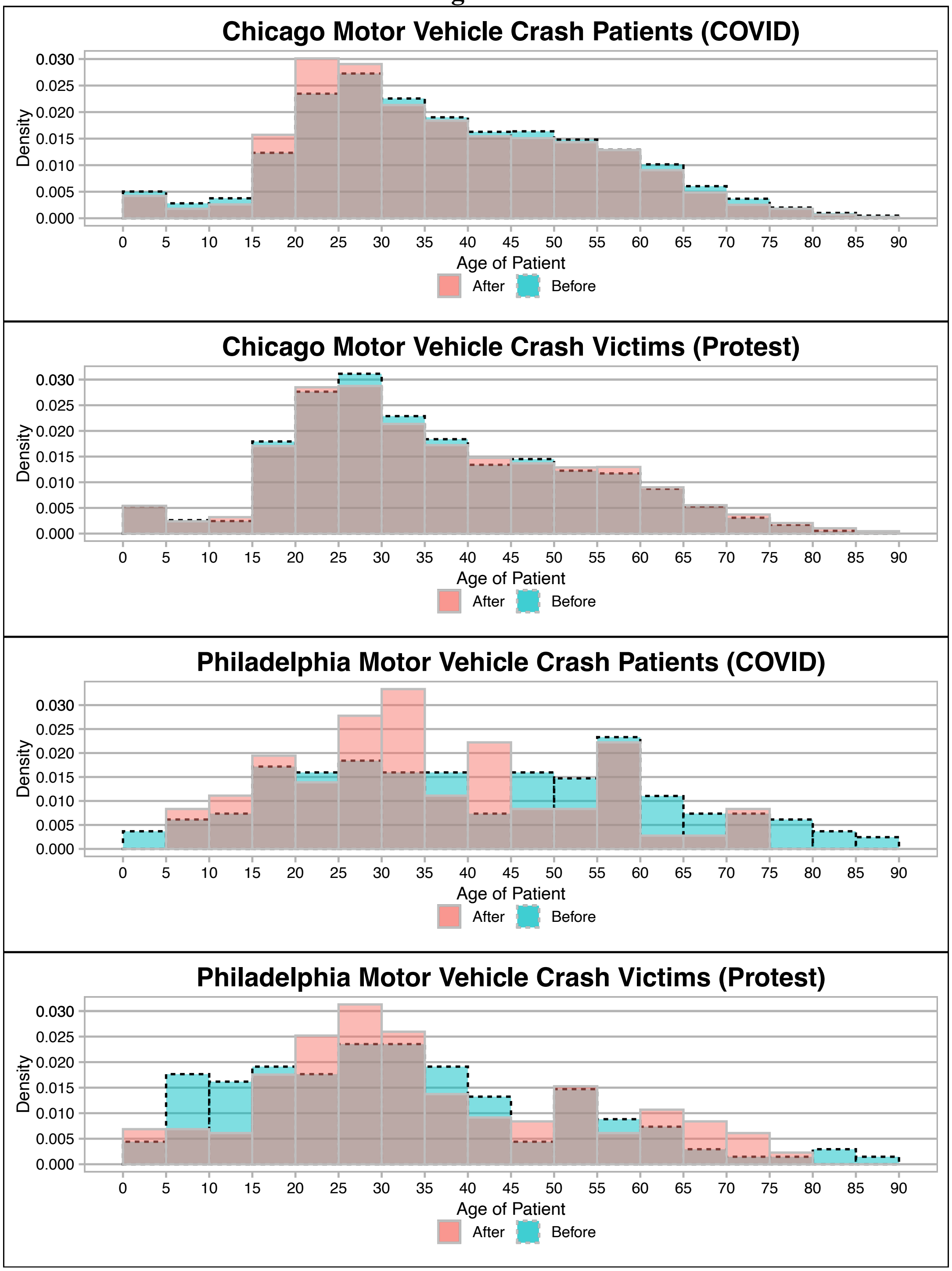
The left axis indicates the scale for the average number of stops conducted by officer pairs who conduct at least one stop on a given day, shown in black. Panel A shows the data for Chicago while Panel B shows the data for Philadelphia. The red vertical line indicates the Pandemic onset while the green 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 8



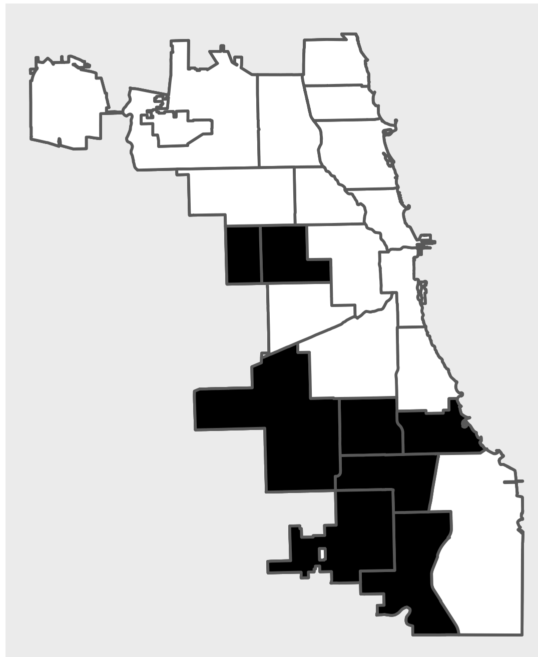
This figure shows the change in mobility in Chicago and Philadelphia over 2020 relative to a baseline established Jan 3 – Feb 6, 2020. Mobility is reported for the Retail and Recreation (blue solid line) and Transit (red dashed line) categories reported by the Google Community Mobility Reports. The red 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 9



The density histograms show 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 10



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	121	81	208	78	990	490
Frisks per Day	57	29	23	22	113	47	111	79
% Male	88%	87%	86%	88%	82%	83%	72%	75%
% Black	65%	64%	70%	78%	64%	64%	74%	80%
Age	35.7	34.5	34.0	33.4	28.3	26.7	34.9	33.0
Frisked/Searched	25%	33%	19%	28%	55%	60%	11%	16%
Contraband|Frisk/Search=1	13.3%	14.9%	12.1%	10.3%	23.9%	23.7%	19.7%	18.8%
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%	11.9%	13.7%	24.5%	27.9%	22.3%	19.7%
Gun|Frisk/Search=1	2.4%	4.3%	3.1%	2.9%	1.3%	2.0%	3.3%	5.2%
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%	4.3%	0.6%	1.8%	3.0%	5.5%
Panel B: Protests								
Stops per Day	126	71	103	29	138	75	661	139
Frisks per Day	42	26	26	7	84	44	133	25
% Male	90%	87%	85%	83%	86%	85%	78%	77%
% Black	63%	62%	76%	71%	63%	70%	79%	75%
Age	34.4	33.6	32.8	35.7	26.5	27.0	32.1	33.3
Frisked/Searched	33%	36%	25%	23%	61%	59%	20%	18%
Contraband|Frisk/Search=1	14.2%	19.2%	14.0%	12.9%	25.4%	29.2%	22.0%	26.0%
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%	11.4%	21.2%	25.4%	24.2%	28.8%
Gun |Frisk/Search=1	5.8%	9.1%	2.8%	4.7%	2.5%	3.0%	3.9%	5.7%
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%	1.6%	4.3%	1.1%	1.3%	4.0%	6.3%

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 protests 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. Black is defined to include all individuals identified as Black, regardless of ethnicity. 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	0.033	0.034	0.004	0.026	0.047	-0.010	0.015	0.016
	(0.016)	(0.018)	(0.006)	(0.012)	(0.016)	(0.026)	(0.008)	(0.015)
After	-0.005	-0.003	0.004	-0.005	0.003	0.001	0.011	2.00E-03
	(0.01)	(0.008)	(0.003)	(0.004)	(0.008)	(0.009)	(0.004)	(0.004)
Treat	-0.347	-0.843	0.164	-0.561	0.514	0.047	0.420	0.066
	(0.381)	(0.324)	(0.123)	(0.174)	(0.361)	(0.439)	(0.177)	(0.19)
Observations	20,634	18,876	20,634	18,876	31,606	15,114	31,606	15,114
Adjusted R2	0.018	0.011	0.006	0.015	0.019	0.005	0.006	0.003
Mean Y	0.192	0.107	0.016	0.027	0.159	0.107	0.033	0.018

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	0.008	0.025	0.006	0.022	-0.004	-0.016	0.021	-0.002
	(0.014)	(0.015)	(0.005)	(0.008)	(0.011)	(0.017)	(0.006)	(0.009)
After	-0.011	-0.005	-0.001	-0.003	-0.024	0.015	0.001	-2.00E-03
	(0.009)	(0.008)	(0.003)	(0.004)	(0.007)	(0.008)	(0.003)	(0.003)
Treat	-0.205	0.172	-0.068	-0.024	-1.047	0.923	0.174	-0.074
	(0.318)	(0.295)	(0.088)	(0.129)	(0.293)	(0.35)	(0.14)	(0.134)
Observations	27,237	20,654	27,237	20,654	42,764	20,682	42,764	20,682
Adjusted R2	0.014	0.012	0.001	0.003	0.016	0.007	0.006	0.003
Mean Y	0.184	0.102	0.013	0.017	0.148	0.092	0.031	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 guns 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. All regressions include police region and year fixed effects as well as a time trend. Robust standard errors clustered at the region level. Data Source: Philadelphia Police Department & Chicago Police Department.

Table 3A: Impact of the Protests 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	0.018	0.021	0.002	0.012	0.007	-0.002	0.003	0.004
	(0.011)	(0.007)	(0.003)	(0.004)	(0.003)	(0.007)	(0.002)	(0.004)
After	0.003	0.002	0.000	0.002	0.0003	-0.002	0.001’	0.000
	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0003)	(0.001)
Treat	0.113	0.033	0.010	0.018	0.099	-0.027	0.043	0.009
	(0.009)	(0.005)	(0.003)	(0.003)	(0.033)	(0.064)	(0.015)	(0.026)
Observations	37,441	69,297	37,441	69,297	369,216	115,040	369,216	115,040
Adjusted R2	0.024	0.010	0.004	0.008	0.007	0.004	0.001	0.001
Mean Y	0.126	0.031	0.009	0.007	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	0.021	0.004	0.009	0.008	0.010	0.005	0.002
	(0.009)	(0.005)	(0.003)	(0.003)	(0.002)	(0.005)	(0.001)	(0.002)
After	-0.003	-0.002	0.0002	-0.001	-0.002’	0.003	0.0002	-0.0003
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0003)	(0.0005)
Treat	0.072	0.009	0.003	0.003	-0.077	0.214	0.022	-0.006
	(0.006)	(0.003)	(0.001)	(0.001)	(0.03)	(0.063)	(0.013)	(0.023)
Observations	47,562	68,056	47,562	68,056	459,404	127,410	459,404	127,410
Adjusted R2	0.015	0.008	0.001	0.002	0.004	0.003	0.001	0.001
Mean Y	0.120	0.033	0.007	0.005	0.014	0.016	0.003	0.002

Note:

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 guns 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. All regressions include police region and year fixed effects as well a time trend. 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	0.016	0.038	0.008	0.017	0.041	0.005	0.016	0.014
	(0.028)	(0.031)	(0.009)	(0.02)	(0.019)	(0.029)	(0.01)	(0.016)
After	0.032	-0.025	-5.00E-03	-0.034	0.003	-0.022	3.00E-03	0.005
	(0.028)	(0.03)	(0.008)	(0.016)	(0.016)	(0.021)	(0.008)	(0.01)
Treat	-0.870	-2.197	0.144	-1.619	0.003	-0.749	-0.008	0.169
	(0.995)	(1.108)	(0.309)	(0.645)	(0.736)	(0.978)	(0.352)	(0.482)
Observations	4,511	2,544	4,511	2,544	9,267	3,217	9,267	3,217
Adjusted R^2	0.030	0.027	0.013	0.010	0.039	0.013	0.011	-0.001
Mean Y	0.247	0.136	0.020	0.045	0.181	0.109	0.034	0.021

Note:

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 guns. 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. All regressions include sector and year fixed effects as well as a time trend. Robust standard errors clustered at the sector level. *Data Source*: Chicago Police Department & Philadelphia 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:

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	0.137	0.184	0.603	-0.023	0.378	0.231	0.084
	(0.043)	(0.061)	(0.212)	(0.175)	(0.147)	(0.071)	(0.095)
After	0.034	-0.016	0.227	-0.074	0.142	-0.068	0.047
	(0.027)	(0.041)	(0.143)	(0.120)	(0.086)	(0.048)	(0.055)
Treat	-1.017	-5.034	9.989	-9.496	-0.347	-6.065	-3.948
	(01.173)	(01.621)	(05.245)	(04.421)	(03.388)	(01.801)	(02.190)
Observations	315	315	315	315	315	315	315
Adjusted R^2	0.612	0.188	0.085	0.085	0.213	0.112	0.385

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	-0.159	-0.035	-0.266
	(0.056)	(0.062)	(0.238)	(0.221)	(0.183)	(0.074)	(0.101)
After	0.019	-0.034	-0.013	-0.132	-0.014	-0.017	-0.014
	(0.022)	(0.035)	(0.143)	(0.131)	(0.120)	(0.047)	(0.057)
Treat	0.57	-1.859	-4.396	-7.428	-4.392	-0.62	-1.056
	(01.083)	(01.539)	(06.670)	(05.960)	(05.157)	(02.170)	(02.568)
Observations	378	378	378	378	378	378	378
Adjusted R^2	0.443	0.095	0.010	0.088	0.086	0.015	0.338

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 and a time trend. 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	0.109	0.293	-0.021	0.056	0.198	-0.138
	(0.045)	(0.067)	(0.223)	(0.210)	(0.162)	(0.079)	(0.112)
After	-0.003	-0.033	0.208	0.025	0.22	-0.117	0.104
	(0.027)	(0.045)	(0.161)	(0.172)	(0.113)	(0.056)	(0.072)
Treat	-1.568	-4.727	4.363	-6.296	-0.992	-6.660	-1.82
	(01.176)	(01.777)	(05.582)	(05.932)	(04.086)	(02.167)	(02.750)
Observations	315	315	315	315	315	315	315
Adjusted R^2	0.398	0.097	0.045	0.009	0.159	0.080	0.220

Panel B	**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	0.060	0.064	0.005	0.039	0.060	0.064	0.005	0.039
	(0.023)	(0.028)	(0.009)	(0.021)	(0.023)	(0.028)	(0.009)	(0.021)
After	0.000	-0.012	0.004	-0.006	0.000	-0.012	0.004	-6.00E-03
	(0.014)	(0.012)	(0.005)	(0.006)	(0.014)	(0.012)	(0.005)	(0.006)
Treat	0.091	-1.189	0.260	-0.801	0.091	-1.189	0.260	-0.801
	(0.526)	(0.474)	(0.187)	(0.264)	(0.526)	(0.474)	(0.187)	(0.264)
Observations	11,180	9,432	11,180	9,432	11,180	9,432	11,180	9,432
Adjusted R^2	0.024	0.016	0.004	0.026	0.024	0.016	0.004	0.026
Mean Y	0.203	0.120	0.020	0.034	0.203	0.120	0.020	0.034

Note:

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 and the regressions in Panel B also include sector fixed effects and a time trend. 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	-0.007	-0.014	0.005	-0.021	-0.045	0.028	-0.029	4.00E-04
	(0.034)	(0.036)	(0.01)	(0.023)	(0.046)	(0.061)	(0.024)	(0.031)
After	0.008	0.010	0.003	0.002	-0.025	-0.018	-0.005	-0.002
	(0.011)	(0.009)	(0.002)	(0.004)	(0.012)	(0.013)	(0.005)	(0.005)
Treat	0.040*	0.024	0.005	0.002	0.031	0.027	0.003	0.003
	(0.019)	(0.02)	(0.005)	(0.009)	(0.024)	(0.035)	(0.01)	(0.013)
Black	-0.027*	-0.014	0.006	-0.010	-0.056	-0.023	0.002	1.00E-03
	(0.012)	(0.01)	(0.003)	(0.005)	(0.011)	(0.012)	(0.005)	(0.005)
After*Treat	0.034	0.042	-0.001	0.038*	0.080	-0.031	0.039	0.013
	(0.029)	(0.029)	(0.007)	(0.016)	(0.042)	(0.054)	(0.021)	(0.025)
After*Black	-0.003	0.008	-0.003	0.008	0.025*	0.024	0.010	0.003
	(0.013)	(0.011)	(0.003)	(0.005)	(0.013)	(0.014)	(0.005)	(0.005)
Treat*Black	0.063	0.041	0.012	0.059	0.040	0.006	0.007	0.012
	(0.024)	(0.025)	(0.007)	(0.014)	(0.026)	(0.039)	(0.011)	(0.015)
Observations	20,634	18,876	20,634	18,876	25,113	15,053	25,113	15,053
Adjusted R2	0.017	0.010	0.006	0.016	0.019	0.005	0.005	0.002
Mean Y	0.192	0.107	0.016	0.027	0.151	0.107	0.031	0.018

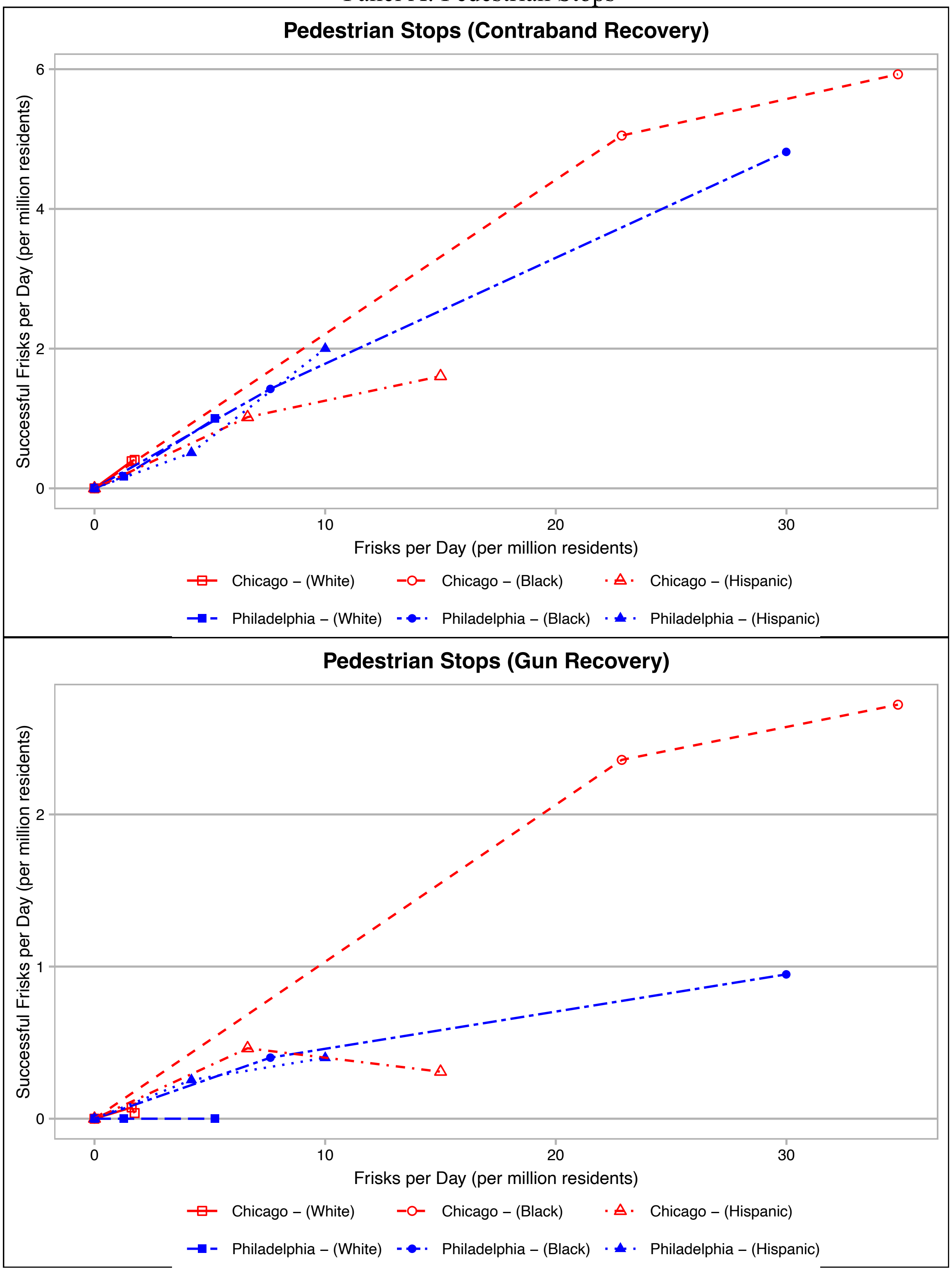
Note:

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 guns. 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, regardless of ethnicity. All regressions include year and police region fixed effects as well as a time trend. Robust standard errors clustered at the region level. *Data Source*: Chicago Police Department & Philadelphia Police Department.

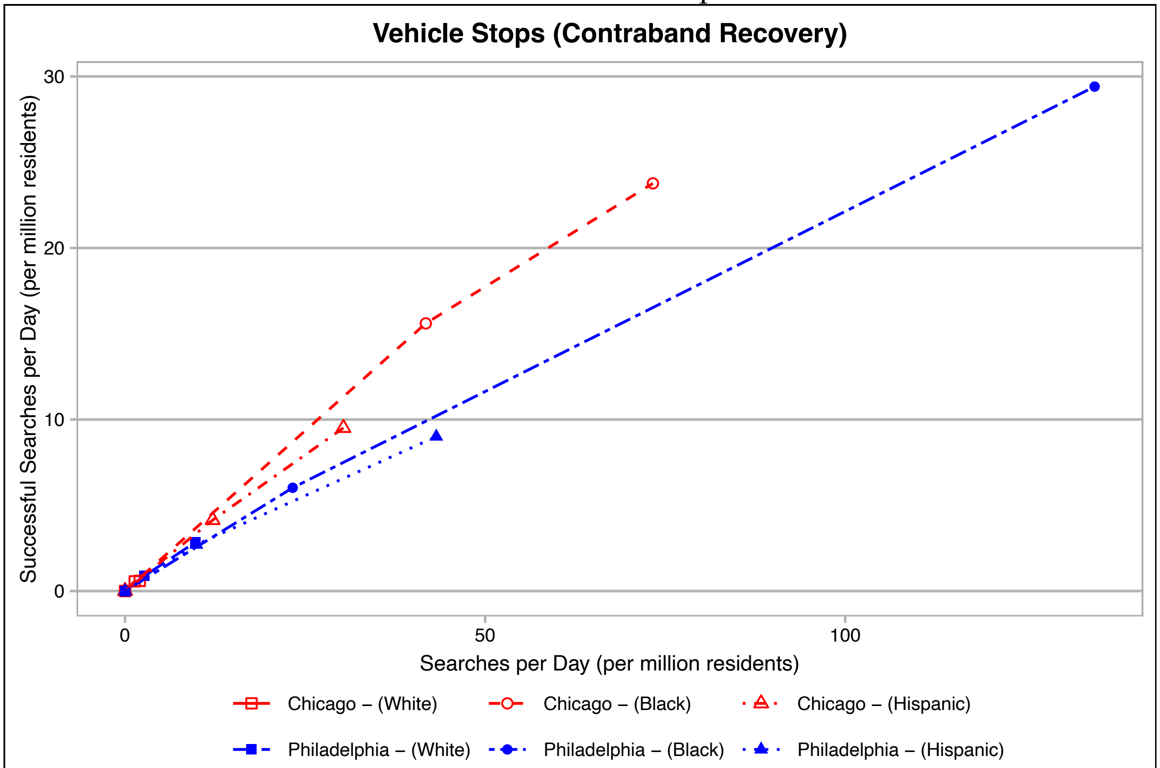
Appendix A

Figure A1. Relationship between Number of Frisks and Successful Frisks, by Race of the Pedestrians/Drivers, and by Contraband and Gun

Panel A: Pedestrian Stops



Panel B: Vehicle Stops



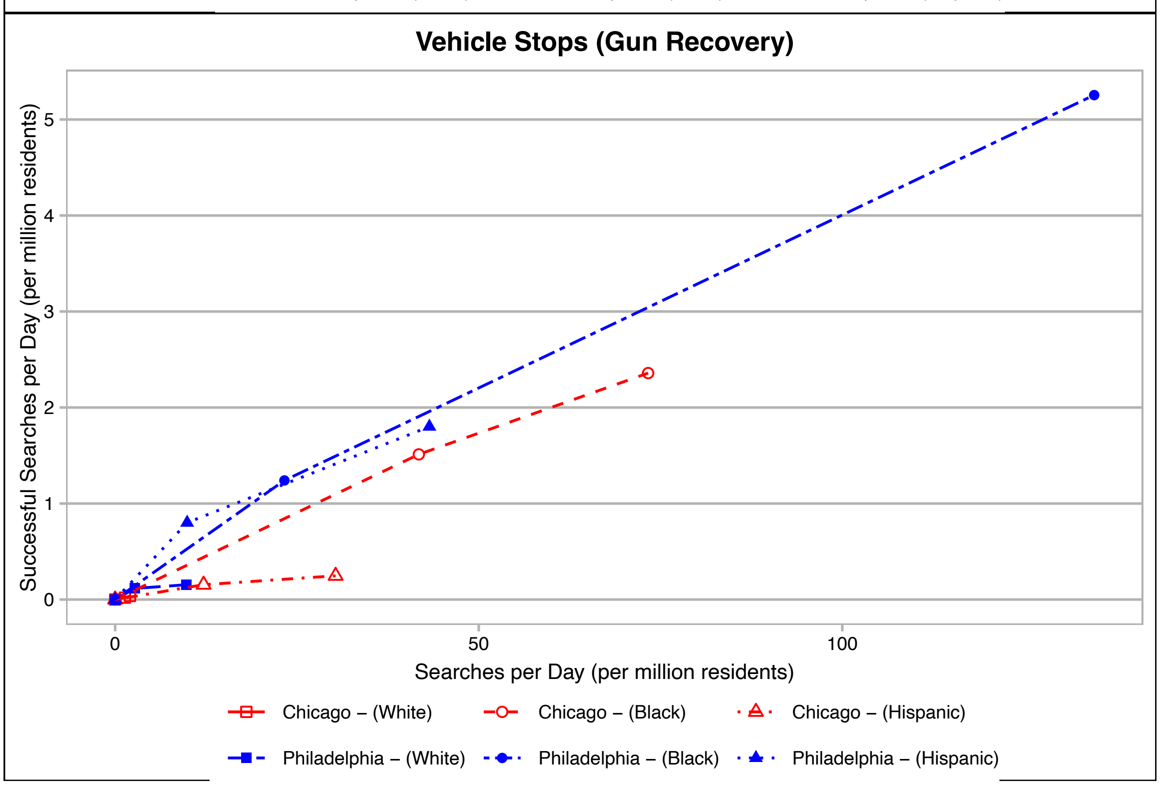
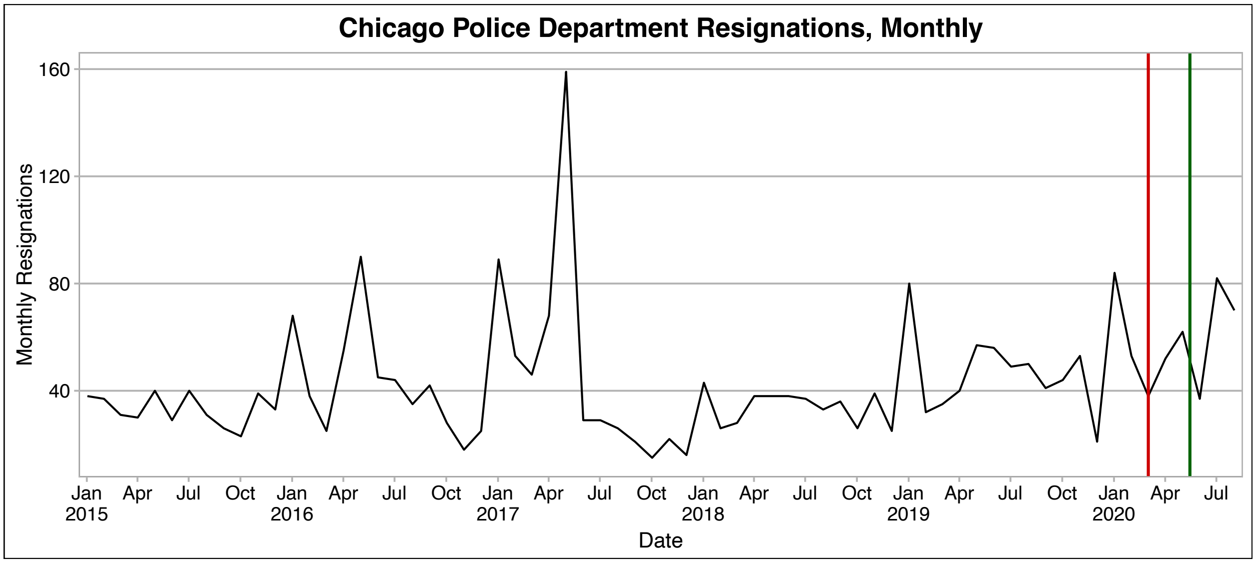


Figure A2.



Shows the monthly number of resignations from the Chicago Police Department between January 2015 – August 2020. Mandatory retirements, upon an officer reaching 63 years of age, are not included. 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: Chicago Police Department

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	-0.574	-0.592	-0.591	-0.437				
	(0.081)	(0.083)	(0.085)	(0.086)				
After	-0.143	-0.052	-0.120	0.010	-1.500	-1.388	-1.690	-1.755
	(0.055)	(0.042)	(0.060)	(0.048)	(0.087)	(0.103)	(0.098)	(0.10)
Treat	-2.646	-4.241	0.284	3.270				
	(2.164)	(1.604)	(2.482)	(1.859)				
Observations	315	315	315	315	63	63	63	63
Adjusted R^2	0.650	0.841	0.540	0.687	0.748	0.689	0.733	0.777

Note:

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	-1.419	-1.299	-1.647	-1.332				
	(0.164)	(0.078)	(0.154)	(0.102)				
After	0.050	0.079	0.003	0.023	-1.463	-1.339	-1.720	-1.419
	(0.039)	(0.042)	(0.075)	(0.035)	(0.166)	(0.078)	(0.152)	(0.097)
Treat	4.851	4.989	4.955	5.262				
	(2.245)	(2.325)	(3.779)	(0.233)				
Observations	378	378	378	378	63	63	63	63
Adjusted R^2	0.871	0.930	0.605	0.829	0.648	0.772	0.661	0.719

Note:

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 on Frisk/Search Hit Rate (incl detainee controls)

Panel A: Protests

	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	0.035	0.032	0.004	0.027	0.043	-0.015	0.017	0.013
	(0.016)	(0.018)	(0.006)	(0.012)	(0.018)	(0.026)	(0.009)	(0.015)
After	-0.004	-0.004	0.004	-0.006	-0.008	0.001	0.003	0.001
	(0.01)	(0.008)	(0.003)	(0.004)	(0.005)	(0.005)	(0.002)	(0.002)
Treat	-0.289	-0.821	0.158	-0.580	-0.069	-0.003	-0.025	-0.011
	(0.382)	(0.325)	(0.124)	(0.176)	(0.038)	(0.037)	(0.018)	(0.015)
Observations	20,551	18,699	20,551	18,699	25,062	15,018	25,062	15,018
Adjusted R2	0.022	0.015	0.006	0.017	0.021	0.005	0.006	0.004
Mean Y	0.192	0.107	0.016	0.027	0.151	0.107	0.031	0.018

Panel B: Pandemic Response

	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	0.010	0.029	0.006	0.021	-0.003	-0.016	0.018	-0.004
	(0.014)	(0.015)	(0.005)	(0.008)	(0.012)	(0.017)	(0.006)	(0.009)
After	-0.012	-0.004	-0.001	-0.003	-0.005	-0.002	-0.003	-2.00E-04
	(0.009)	(0.008)	(0.003)	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)
Treat	-0.290	0.176	-0.062	-0.010	0.003	0.009	-0.041	0.015
	(0.319)	(0.297)	(0.089)	(0.13)	(0.032)	(0.027)	(0.014)	(0.011)
Observations	27,145	20,451	27,145	20,451	34,083	20,555	34,083	20,555
Adjusted R2	0.019	0.015	0.002	0.005	0.018	0.007	0.007	0.004
Mean Y	0.184	0.102	0.013	0.017	0.144	0.092	0.030	0.014

Note:

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 guns 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. All regressions include police region and year fixed effects as well as a time trend. Controls are included for detainee age, race, 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 A3: 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	0.030	0.045	0.005	0.034	0.046	0.004	0.022	0.015
	(0.015)	(0.018)	(0.005)	(0.013)	(0.016)	(0.027)	(0.009)	(0.015)
Observations	3,623	1,945	3,623	1,945	3,012	803	3,012	803
Adjusted R2	0.027	0.029	0.008	0.015	0.048	0.004	0.019	-0.035
Mean Y	0.274	0.170	0.028	0.076	0.231	0.136	0.044	0.034

Note: 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 guns 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. All regressions include police region fixed effects as well as a time trend. Robust standard errors clustered at the region level. Data Source: Chicago Police Department & Philadelphia Police Department.

Table A4: Improvement in Recovery Relative to Random Search/Frisk

	Pedestrian		Vehicle	
	Gains Over Random Frisks	Gains Over Random Frisks (%)	Gains Over Random Searches	Gains Over Random Searches (%)
Chicago (Contraband)	500.319	25.80%	771.33	13.01%
Philadelphia (Contraband)	5.14	1.21%	492.588	18.62%
Chicago (Gun)	299.862	35.56%	76.805	15.80%
Philadelphia (Gun)	49.289	40.43%	172.481	31.98%

Note: This table reports the additional "all contraband" and gun recovery from the search/frisk behavior exhibited by officers in the data relative to what would occur under random search/frisk behavior. Total recoveries are annualized assuming that the daily number of frisks/searches would remain at the average level observed in the 6 weeks after the start of the George Floyd protests in each city.

Table A5: Impact on Frisk/Search Hit Rate (including officer fixed effects)

Panel A: Protests

	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	0.032	0.041	-0.003	0.007	0.055	0.039	0.018	0.006
	(0.027)	(0.03)	(0.008)	(0.017)	(0.03)	(0.044)	(0.02)	(0.024)
After	0.008	0.008	0.002	-0.002	-0.004	0.002	0.004	1.00E-03
	(0.017)	(0.012)	(0.005)	(0.006)	(0.009)	(0.007)	(0.005)	(0.003)
Treat	-0.054	-0.179	0.077	-0.439	-0.117	-0.042	0.009	0.011
	(0.557)	(0.525)	(0.235)	(0.255)	(0.093)	(0.075)	(0.035)	(0.025)
Observations	20,634	18,876	20,634	18,876	31,606	15,114	31,606	15,114
Adjusted R2	0.166	0.122	0.108	0.140	0.282	0.151	0.209	0.133
Mean Y	0.192	0.107	0.016	0.027	0.159	0.107	0.033	0.018

Panel B: Pandemic Response

	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	0.023	-0.004	0.006	0.029	0.014	-0.038	0.028	-0.009
	(0.017)	(0.028)	(0.008)	(0.013)	(0.022)	(0.018)	(0.01)	(0.011)
After	-0.010	-0.005	0.000	-0.009	-0.012	-0.007	-0.006	-2.00E-03
	(0.014)	(0.012)	(0.004)	(0.006)	(0.006)	(0.005)	(0.003)	(0.002)
Treat	0.134	0.322	0.047	-0.137	-0.067	-0.009	-0.089	0.021
	(0.511)	(0.469)	(0.121)	(0.188)	(0.043)	(0.047)	(0.032)	(0.021)
Observations	27,237	20,654	27,237	20,654	42,764	20,682	42,764	20,682
Adjusted R2	0.149	0.118	0.125	0.171	0.258	0.167	0.176	0.108
Mean Y	0.184	0.102	0.013	0.017	0.148	0.092	0.031	0.014

Note:

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 guns 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. All regressions include officer, PSA and year fixed effects as well as a time trend. Robust standard errors clustered at the region level. Data Source: Philadelphia Police Department & Chicago Police Department.

Table A6: Impact on Frisk/Search Hit Rate with Constant Geographic Distribution

Panel A: Protests

	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	0.010	0.039	-0.002	0.026	0.048	0.002	0.019	-0.001
	(0.016)	(0.018)	(0.006)	(0.012)	(0.016)	(0.026)	(0.009)	(0.012)
After	-0.004	-0.003	0.005	-0.006	0.004	0.001	0.011	2.00E-03
	(0.01)	(0.008)	(0.003)	(0.004)	(0.008)	(0.009)	(0.004)	(0.004)
Treat	-0.279	-0.844	0.217	-0.591	0.538	0.018	0.372	0.083
	(0.379)	(0.325)	(0.117)	(0.178)	(0.361)	(0.439)	(0.178)	(0.19)
Observations	20,633	18,876	20,633	18,876	31,590	15,114	31,590	15,114
Adjusted R2	0.016	0.012	0.006	0.017	0.020	0.006	0.007	0.002
Mean Y	0.189	0.107	0.016	0.028	0.159	0.107	0.033	0.018

Panel B: Pandemic Response

	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	0.016	0.010	0.017	-0.004	-0.030	0.027	-0.003
	(0.014)	(0.015)	(0.005)	(0.008)	(0.011)	(0.016)	(0.006)	(0.009)
After	-0.010	-0.006	-0.001	-0.003	-0.025	0.016	0.001	-2.00E-03
	(0.009)	(0.008)	(0.003)	(0.004)	(0.007)	(0.008)	(0.003)	(0.003)
Treat	-0.174	0.134	-0.065	-0.035	-1.145	0.962	0.166	-0.066
	(0.318)	(0.295)	(0.089)	(0.129)	(0.293)	(0.35)	(0.14)	(0.134)
Observations	27,239	20,653	27,239	20,653	42,755	20,674	42,755	20,674
Adjusted R2	0.014	0.012	0.002	0.004	0.016	0.007	0.006	0.003
Mean Y	0.184	0.103	0.013	0.017	0.149	0.092	0.031	0.014

Note:

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 guns 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. All regressions include officer, PSA and year fixed effects as well as a time trend. The stops in the after*treat period are sampled (with replacement) from the raw data to keep the distribution of stops across police districts equal to that in the before period of 2020. Robust standard errors clustered at the region level. Data Source: Philadelphia Police Department & Chicago Police Department.

Table A7: 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

Note: 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 A8: 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	0.042	0.032	-0.001	0.015	0.044	0.033	0.001	0.026	0.041	0.043	0.002	0.028
	(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	-0.001	0.013	0.002	0.007	-0.002	-0.012	0.003	-0.006	0.001	-0.025	0.004	-0.002
	(0.007)	(0.006)	(0.002)	(0.003)	(0.012)	(0.01)	(0.004)	(0.005)	(0.014)	(0.011)	(0.005)	(0.006)
Treat	-0.162	-0.164	0.054	0.001	-0.252	-1.095	0.080	-0.501	-0.162	-1.407	0.120	-0.312
	(0.136)	(0.109)	(0.044)	(0.058)	(0.405)	(0.336)	(0.13)	(0.174)	(0.411)	(0.33)	(0.132)	(0.176)
Observations	34,227	31,987	34,227	31,987	20,141	18,354	20,141	18,354	19,995	18,132	19,995	18,132
Adjusted R^2	0.022	0.014	0.005	0.013	0.019	0.012	0.005	0.016	0.019	0.015	0.005	0.016
Mean Y	0.200	0.115	0.016	0.029	0.195	0.109	0.016	0.028	0.197	0.110	0.016	0.028

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	0.039	-0.002	0.034	0.027	0.069	-0.019	0.024	0.015	0.064	-0.018	0.038	0.011
	(0.012)	(0.02)	(0.006)	(0.011)	(0.017)	(0.027)	(0.009)	(0.016)	(0.016)	(0.027)	(0.009)	(0.016)
After	-0.005	0.001	0.003	-0.001	0.001	-0.012	0.004	0.000	0.010	0.005	0.005	0.004
	(0.005)	(0.007)	(0.003)	(0.003)	(0.009)	(0.011)	(0.004)	(0.005)	(0.01)	(0.012)	(0.005)	(0.005)
Treat	-0.017	0.059	-0.074	-0.042	0.452	-0.692	0.056	-0.095	0.815	0.132	0.134	0.066
	(0.123)	(0.148)	(0.065)	(0.068)	(0.362)	(0.448)	(0.182)	(0.19)	(0.363)	(0.445)	(0.185)	(0.193)
Observations	54,178	25,674	54,178	25,674	31,712	14,995	31,712	14,995	31,619	14,937	31,619	14,937
Adjusted R^2	0.017	0.006	0.006	0.004	0.020	0.006	0.006	0.003	0.019	0.008	0.007	0.002
Mean Y	0.159	0.106	0.036	0.019	0.162	0.107	0.034	0.017	0.161	0.106	0.035	0.018

Note:

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 guns. 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 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 (2016-2020 for Chicago and 2015-2020 for Philadelphia). After = 1 beginning on 29 May for Chicago and 3 June for Philadelphia and 0 otherwise; Treat=1 for 2020 and 0 otherwise. All regressions include sector and year fixed effects as well as a time trend. Robust standard errors clustered at the sector level. Data Source: Chicago Police Department and Philadelphia Police Department.