IntegrAssure

OFFICE OF THE INDEPENDENT CONSENT DECREE MONITOR FOR THE CITY OF AURORA

MONITOR'S INTRODUCTION TO THE NPI REPORT:

ENFORCEMENT DATA ANALYSIS FOR THE AURORA, COLORADO POLICE DEPARTMENT

December 18, 2023

TABLE OF CONTENTS

INTRODUCTION

As part of our role as Independent Consent Decree Monitor, we are releasing a Report by NPI entitled "Enforcement Data Analysis for the Aurora Colorado Police Department" (the "Report"). The Report provides an in-depth analysis of law enforcement activities in Aurora over a six-year period, offering insights into the trends and dynamics of policing in the community.

The Report was authored by a research team from the National Policing Institute (NPI), led by Dr. Robin Engel, a nationally recognized leader in criminal justice and police reform. The NPI team is known for their evidence-based approach to policing, emphasizing the importance of data-driven strategies in enhancing public safety and community trust. NPI's commitment to advancing effective, just, and equitable policing aligns with the values of the Consent Decree.

The Report examines several key aspects of policing activities in Aurora, including trends in criminal incidents, criminal summonses, arrests, and uses of force, from January 1, 2017, to December 31, 2022.

The Report provides insights into disparities that exist in enforcement activity and potential gaps in policies and procedures and how data should be collected moving forward.

METHODOLOGY AND INHERENT LIMITATIONS OF THE NPI ANALYSES

The Report fully details the methodologies that were employed by NPI in its analyses and the inherent strengths and limitations of each methodology. Of note are the following:

- In order to analyze whether APD arrests people or uses force against them at different rates depending on their race/ethnicity, the NPI Report in part used "benchmark comparisons." This means that NPI compared (1) the percent of people arrested, or against whom force is used, who are of a certain race/ethnicity to (2) the percent of people in a benchmark comparison group who are of that same race/ethnicity.[1] Each benchmark comparison group yields a different outcome relative to disparities between racial/ethnic groups.[2]

- Although both census and non-census benchmarks are included in the Report, NPI believes that "non-census benchmarks," i.e., benchmark comparison groups that do not rely on the census population, are more meaningful because they better approximate the population of individuals who are "at risk" of enforcement action.[3] NPI used two kinds of

[1] *See* NPI Report pg. 21-22, 54.
[2] *See* NPI Report pg. 21.
[3] *See* NPI Report pg. 22.

non-census benchmarks: (1) people reported to APD as being criminal suspects, and (2) people who were arrested or issued criminal summonses.[4]

- NPI notes, that if bias affects who becomes a part of a benchmark comparison group to begin with, this may cause the benchmarking analysis to underestimate the amount of racial/ethnic disparity in police enforcement actions.

- Similarly, NPI notes that because criminal suspect data depends on choices made by members of the public regarding who to report as potential criminal suspects and which crimes to report, "reported crimes may themselves be biased against offenders of certain racial/ethnic groups."[5]

- NPI also conducted multivariate regression analysis to predict to what extent race/ethnicity might influence whether APD uses force.[6] This analysis, by design, utilized data for all arrested individuals with the express purpose of making predictions as to how likely it is that force was used against an individual of a particular race or ethnicity who is arrested.[7] The Report did not analyze how likely it is that force is going to be used against an individual who is <u>not</u> arrested.

Most notably, with respect to the presence of bias or racial profiling, the Report points out that limitations arise from both the nature of the data available and the complexities inherent in policing and social interactions. While the report is able to measure racial disparities (as opposed to bias or racial profiling) and indicates small to marginal disparities in arrests and uses of force for Black individuals compared to White individuals, any level of disparity is a matter of concern. Specifically, the Report cautions:

> *It is important to note two caveats to the findings presented in this report. First, no statistical analysis using these data can determine if APD officers engage in racially biased enforcement actions. While it is possible to estimate racial disparities in enforcement actions (i.e., differences in outcomes across racial/ethnic groups) using a combination of statistical analyses, it is beyond the scope of any quantitative analysis to determine if any disparities observed are due to officer bias or discrimination.*
>
> *Second, no single analysis can determine definitively if APD enforcement actions are racially disparate. Each type of statistical analysis has strengths and limitations that should be considered when interpreting the findings. It is possible that analyses employing different techniques or data sources produce conflicting findings. The purpose of conducting multiple analyses using a variety*

[4] *See* NPI Report pp. 4-8, 21-22, 48, 55.
[5] *See* NPI Report pg. 21-23.
[6] *See* NPI Report pg. 67.
[7] *See* NPI Report pg. 62-63.

of data sources is to develop a more comprehensive understanding of APD enforcement patterns.[8]

In light of all these limitations, it is essential that findings from the Report are interpreted with caution. While they provide valuable insights, they represent a piece of a much larger puzzle. A comprehensive assessment of bias or racial profiling in law enforcement requires a multi-faceted approach that includes qualitative research, community engagement, and ongoing, transparent dialogue. It is through this broader lens that we can begin to more fully understand and address the complex issues of bias and racial profiling in policing.

NOTABLE FINDINGS

The report contains several notable findings:

- Decrease in Criminal Summons Issued: The report notes a consistent decline in the number of criminal summonses issued by the Aurora Police Department (APD) over the six-year period. From a peak of over 5,000 in 2017, there was a significant drop to around 2,300 in 2022. This decline amounted to 54.1% over six years.

- Stable or Reduced Racial Disparities in Criminal Summons: The racial and ethnic distribution of individuals who received criminal summonses remained consistent from 2017 to 2022. Despite the overall reduction in summonses, the distribution across racial and ethnic groups did not show significant disparities. This stability suggests that the decrease in summonses was applied uniformly across different demographics.

- Significant Decline in Arrests Post-COVID: The Report highlights a substantial reduction in APD arrests following the onset of the COVID-19 pandemic. There was a nearly 47% decline in overall arrests in the post-2020 period compared to the pre-2020 period. This decrease was consistent across all racial and ethnic groups and was more pronounced for less serious offenses.

- Reduced Racial Disparities in Arrests: Using its non-census benchmarks, the Report found small to marginal racial/ethnic disparities in arrests, with post-COVID disparities decreasing and, in some benchmarks, showing that Black and Hispanic individuals were less likely to be arrested compared to White individuals.

- Increased Proportion of Arrests for Serious Offenses Post-COVID: Although there was a decline in the total number of arrests post-COVID, the proportion of arrests for more serious and violent offenses increased.

- Stability in Use of Force Incidents: The number of individuals against whom force was used by the APD remained relatively stable throughout the six-year period. This stability,

[8] *See* NPI Report pg. ix.

coupled with the significant decrease in arrests, resulted in an increased percentage of arrestees experiencing use of force. However, the overall number of uses of force incidents did not increase significantly.

- No Significant Disruptions in Use of Force Post-COVID: Unlike the trends observed in criminal summonses and arrests, the use of force by APD officers did not experience significant disruptions due to the onset of the COVID-19 pandemic or other seminal events during the study period.

- Significant Racial Disparities in Use of Force Continue to Exist: Looking at the population of arrestees, while Black arrestees are significantly more likely to have force used against them compared to White arrestees (after controlling for situational, legal and arrestee characteristics), there are no statistical differences in use of force against Hispanic arrestees compared to White arrestees. Racial differences were reduced post-COVID compared to the prior three years.

- Increase in Serious Violent Crime: From 2017 to 2022, the city of Aurora experienced a 44% increase in Part I crimes (serious crimes such as murder, rape, robbery, aggravated assault, burglary, larceny, auto theft, and arson) and an 82% rise in violent crime. This persistent upward trend of reported crime was not significantly altered by external events like the COVID-19 pandemic, indicating a growing concern for public safety.

IMPACT OF COVID-19 AND OTHER SEMINAL EVENTS

While the onset of the COVID-19 pandemic in March 2020 was found to mark a significant turning point in some data trends, the COVID emergency coincides and overlaps with several other seminal events that potentially influenced law enforcement activities and crime patterns. The Report acknowledges the complexities in attributing specific changes to these individual events.

Among these other post-COVID seminal events that <u>may</u> have had an effect on law enforcement activity and crime patterns are the following:

- Officer-involved death of George Floyd in Minneapolis – May 2020
- Enactment of Colorado SB 20-217: Enhance Law Enforcement Integrity – July 2020
- AG launches pattern or practice investigation – August 2020
- Independent Review Panel report released - February 2021
- Indictment of officers involved in McClain death - September 2021
- City enters into Consent Decree - November 2021
- Monitor Selected and Monitorship Begins – February 2022
- APD Chief Vanessa Wilson terminated - April 2022
- Interim APD Chief Dan Oates hired - June 2022

While It is clear that the significant shifts in crime and law enforcement activities started with the onset of COVID, teasing out the specific impact of the remaining seminal events beyond COVID on the observed changes presents a complex challenge for several reasons:

- A number of the seminal events occurred in close temporal proximity to each other and to the onset of the COVID-19 pandemic. This makes it difficult to isolate the effects of individual events on law enforcement and crime trends.

- Each event had its own set of implications, and their effects are likely to be multifaceted and interconnected. This complexity adds to the challenge of attributing specific changes in crime and enforcement patterns to individual events.

- The data used in the report may not have the granularity or the specific variables necessary to directly link changes in crime and enforcement activities to particular seminal events. Without detailed, event-specific data, drawing direct causal links remains speculative.

- A number of the seminal events, along with the COVID-19 pandemic, had broad socioeconomic and psychological impacts on the community. These broader effects could indirectly influence crime rates and police activity, further complicating the analysis.

Given these challenges, it is not feasible to definitively attribute the changes observed in the report to specific seminal events. However, acknowledging the presence and potential influence of these events is important in understanding the broader context of the observed trends.

CONCLUSION

In conclusion, the Report offers a multifaceted, but nonetheless limited, view of law enforcement in Aurora. It highlights trends in the areas which were analyzed over the period of time examined. As we move forward, we will utilize this information, along with other information coming from a variety of other sources, in determining the progress that Aurora is making in its reform efforts pursuant to the Consent Decree.

APPENDIX A: ENFORCEMENT DATA ANALYSIS FOR THE AURORA COLORADO POLICE DEPARTMENT (NPI)

[The remainder of this page is intentionally blank. The NPI Report follows on the next page.]

National Policing Institute
PURSUING EXCELLENCE THROUGH SCIENCE AND INNOVATION

ENFORCEMENT DATA ANALYSIS FOR THE AURORA, COLORADO POLICE DEPARTMENT

PREPARED FOR:

IntegrAssure, LLC., the City of Aurora, and the Aurora Police Department

PREPARED BY:

Robin S. Engel, Ph.D.

Jennifer C. Cherkauskas, Ph.D.

Nicholas Corsaro, Ph.D.

Murat Yildirim, M.S.

Hannah McManus, Ph.D.

Ryan Fisher, Ph.D.

DECEMBER 2023

The findings and recommendations presented within this report are from the authors and do not necessarily reflect the official positions or opinions of IntegrAssure, LLC., the City of Aurora, or the Aurora Police Department. Please direct all correspondence regarding this report to Robin S. Engel, Ph.D., Senior Vice President, National Policing Institute, 2550 S. Clark Street, Suite 1130, Arlington, VA 22202; 202-833-1254; rengel@policinginstitute.org.

About the National Policing Institute

Established in 1970, the National Policing Institute (NPI, formerly the National Police Foundation) is the oldest nationally known 501(c)(3) nonprofit, nonpartisan, independent research organization dedicated to improving policing in the United States. The National Policing Institute supports change-makers in policing, communities, and government by harnessing the power of science and innovation to promote public safety for all. The National Policing Institute operates with independence and objectivity. Our work identifies ways to improve policing, ignite a spirit of collaboration among officers and the communities they serve, and use rigorous scientific study results to address the most complex public safety issues facing neighborhoods, cities and towns, states, and the nation. Over the last 53 years, the National Policing Institute's work has remained a catalyst for significant change in policing and communities, contributes to scholastic exploration and discovery, informs policymakers, community members, and practitioners alike, and serves as a model for the systematic and fact-based examination of real-world challenges. To accomplish this mission— Pursuing Excellence through Science and Innovation—the National Policing Institute works closely with those working in and affected by policing across the United States and internationally. Today, the National Policing Institute continues to advance the principles of 21st-century democratic policing through its work. Though many may have ideas worthy of consideration, the National Policing Institute offers actionable solutions to the challenges confronting communities and policing leaders.

About the Authors

Robin S. Engel, Ph.D., serves as Senior Vice President at the National Policing Institute, following over 25 years in academic positions within higher education institutions. As an award-winning researcher, she has partnered with dozens of police agencies in the U.S. and internationally, served as Principal Investigator for over 100 research studies and projects, and ranked among the top academics nationally in criminal justice/criminology. From 2015 – 2019, she served as Vice President for Safety and Reform at the University of Cincinnati, where her executive duties included oversight of daily operations and successful implementation of comprehensive police reforms in the aftermath of a fatal police shooting of an unarmed motorist.

Dr. Engel's partnerships with law enforcement agencies have been honored with the 2008 IACP/Motorola Webber Seavey Award for Excellence in Law Enforcement, the 2009 IACP/West Award for Excellence in Criminal Investigations, and the 2008 National Criminal Justice Association's Outstanding Criminal Justice Program Award. She received the Distinguished Alumni Award from the University at Albany in 2017, and the 2022 O.W. Wilson Award from the Academy of Criminal Justice Sciences. As a top scholar and trusted leading authority in police science, Dr. Engel oversees the Institute's vast scientific investments designed to translate and integrate research into practice.

Dr. Engel has led traffic/pedestrian stop and use of force studies for the Pennsylvania State Police, Ohio State Highway Patrol, Arizona Department of Public Safety, Nebraska State Patrol, State of Vermont, and American Civil Liberties Union, along with local police departments in Cincinnati (OH), Seattle (WA), Tulsa (OK), Oklahoma City (OK), Colorado Springs (CO), and Fairfax County (VA). She has also published extensively on the topics of traffic stop and use of force data analyses within the top academic peer-reviewed journals in the field, has served as an expert on invited panels and conferences, and provided expert testimony.

Jennifer Calnon Cherkauskas, Ph.D., is a Director at the National Policing Institute. Dr. Cherkauskas previously worked as a Senior Research Associate and a Research Consultant at the International Association of Chiefs of Police (IACP)/University of Cincinnati (UC) Center for Police Research and Policy. Over the last twenty years, she has worked with state, local, and campus police agencies on a variety of research and technical assistance projects. She currently works with police agencies across the country as part of multiple research projects examining police use of force, traffic stops, and training evaluations. Dr. Cherkauskas earned her Ph.D. in Crime, Law, and Justice from The Pennsylvania State University.

Nicholas Corsaro, Ph.D., serves as a statistical consultant to the NPI team. He is an Associate Professor of Criminal Justice at the University of Cincinnati. Dr. Corsaro has served as a principal investigator for multiple contracts and grants totaling over seven million dollars with a focus on improving police practice. Dr. Corsaro's research focuses extensively on quantitative analysis of data used in police field settings. His scholarship appears in a diverse range of highly ranked scholarly outlets in the fields of criminology, public health, and evaluation research. He received his doctorate from the School of Criminal Justice at Michigan State University.

Murat Yildirim, M.S., joined the National Policing Institute as a Research Data Manager in January 2023. Previously, he worked as a Research Associate for the IACP/UC Center for Police Research and Policy. He graduated from the National Police Academy of Turkey in 1997 with a B.A. and earned his Master of Science in Forensic Science from the University of Ankara in Ankara, Turkey, in 2002. He worked as a senior police officer and taught classes and courses at the National Police Academy of Turkey between 1997 and 2011. His work includes the acquisition, collection, cleaning/preparation, analysis, visualization, and mapping of crime data.

Hannah D. McManus, Ph.D., is a Director at the National Policing Institute. Prior to joining NPI, she worked for five years as a Research Associate for the IACP/UC Center for Police Research and Policy. Dr. McManus participates in research and evaluation in the criminal justice field, with a specific focus on police practice. Her current work includes research on the development and implementation of training for police, best practices in crisis response, and the examination of community-police relations. As an advocate for academic-practitioner partnerships to promote evidence-informed and best practices, Dr. McManus has worked directly with law enforcement agencies across the country. She currently serves as Co-Principal Investigator on multiple federally awarded grants and has published in both peer-reviewed and practitioner-focused outlets, earning several awards for those publications. Dr. McManus earned her doctorate from the School of Criminal Justice at the University of Cincinnati.

Ryan Fisher, Ph.D., is a Senior Research Associate at the National Policing Institute. He joined the National Policing Institute in August of 2023. Previously, Dr. Fisher held roles at the American Institutes for Research and the City of Philadelphia, providing research and technical assistance on projects related to policing, crime and violence prevention, and criminal justice reform. Dr. Fisher's background is in quantitative and spatial analysis, and his published research includes work on policing practices, environmental criminology, and quantitative methods. In addition, he has extensive experience in public policy work, including researcher-practitioner partnerships and collaboration with local and federal criminal justice agencies. He earned his Ph.D. in Criminology and Criminal Justice from Temple University.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

On November 16, 2021, the Colorado Office of the Attorney General announced its initiation of a consent decree with the City of Aurora, Colorado that mandated oversight of the Aurora Police Department (APD), Aurora Fire Rescue, and Aurora Civil Service Commission. It was specifically recommended by an investigative team appointed by Colorado's Attorney General, Phil Weiser, to conduct a pattern or practice investigation. On February 14, 2022, IntegrAssure, LLC., was appointed as the Independent Consent Decree Monitor to oversee these agencies' implementation of consent decree mandates and ensure progression toward compliance goals.

IntegrAssure engaged the National Policing Institute (NPI) to support the development of baseline measures that may be used to examine changes in police enforcement actions as the APD implements efforts to meet consent decree requirements. This work will facilitate IntegrAssure's assessment of changes in the APD's engagement with community members, including changes to racial disparities in officers' interactions, arrests, and uses of force in the community over time.

The data collection and analytic strategy for this work was guided by six research questions pertaining to the enforcement activities of the APD over time and across groups of community members. These questions included:

1. What are the trends and patterns in APD's criminal summonses, arrests, and uses of force over time?
2. Does the frequency of criminal summonses, arrests, and uses of force shift significantly after seminal events?
3. Do rates of arrest and use of force experienced by different racial/ethnic groups align with their representation among the populations at risk of experiencing enforcement actions by the APD?
4. What factors predict the likelihood of use of force by APD officers?
5. Are community members' race, ethnicity, or gender associated with the type or severity of force used by APD officers?
6. What factors predict the likelihood of injuries to community members or officers during use of force incidents?

DATA SOURCES

Four data sources were examined to identify trends in APD's enforcement activities over a six-year period (Jan 1, 2017 – Dec 31, 2022), including (1) reported criminal offenses, (2) criminal summonses, (3) arrests, and (4) uses of force. The collection and management process for three of the four data sources (reported criminal offenses, criminal summonses, and arrests) were conducted with efficacy and provide an acceptable level of confidence in the reliability and validity of the data and subsequent analyses. However, significant data limitations and quality concerns with the APD's use of force reporting constrained the NPI team's capacity to compile and analyze use of force data to support this study. In particular, the inability to link officers and subjects in use of force incidents and the prevalence of missing data on use of force reports prohibited many of the analyses initially planned by the NPI team. As a result, research questions 5 and 6 could not be answered.

Data Provided by APD to Support Analyses (2017–2022)	
Reported Criminal Offenses	• Data aggregated to incident (n = 33,495 incidents) and suspect (n = 35,889 suspects) levels. • Used to examine trends in criminal incidents, provide context to APD enforcement activities, and facilitate benchmark comparisons to known criminal suspects.
Criminal Summonses	• Charge-level data aggregated to the individual level (n = 20,922 individuals). Excludes traffic summons (without a criminal charge). • Used to examine trends in criminal summonses and facilitate benchmark comparisons for those who experienced use of force by APD officers.
Arrests	• Data aggregated to the individual arrestee level (n = 44,954 arrestees). • Used to examine arrest trends, examine racial/ethnic disparities, create benchmark comparisons for individuals who experience use of force by APD officers, and identify factors that predict use of force against arrestees.
Use of Force	• Data analyzed at the subject level (i.e., individual who experienced use of force in a single incident) (n = 3,783 subjects; across 3,518 use of force incidents). • Used to examine trends in APD officers' use of force and facilitate benchmark comparisons using force data as the numerator. *__Note__: Use of force data were extracted for the NPI team at multiple units of analysis (e.g., incident, officer, subject, force action), complicating data aggregation to the subject level. Inconsistencies in these data limited the analysis of APD's use of force data.

STATISTICAL ANALYSES

The purpose of this report is to develop baseline measures in APD enforcement patterns to assess the agency's progress over time as reform efforts are implemented. Based on the APD enforcement data available, four types of statistical analyses are conducted to examine enforcement patterns and trends: (1) descriptive analyses, (2) time series analyses, (3) benchmark comparisons, and (4) multivariate analyses.

Descriptive analyses summarize and present outcome count data. They provide a basis for understanding basic patterns and distributions in the data and offer an initial assessment of the general trends for a single variable or the potential correlations between two variables.

Time series analyses consider how patterns and trends in police enforcement actions fluctuate over time. Across the six-year period examined, numerous seminal events occurred that may have impacted – or disrupted – preexisting patterns in crime and police enforcement. The NPI team identified ten seminal events to consider when analyzing trends in crime and APD enforcement activities over time.[1] The impact of these events on APD enforcement activities is assessed using interrupted time series analyses.

Benchmark analysis is a statistical method used to examine and assess potential disparities in outcomes across racial/ethnic groups using a reference point (or benchmark) against which rates for different groups can be compared. This analysis relies on the availability of reliable and valid benchmark comparisons. To examine disparities in APD police enforcement activity, the benchmark population should accurately estimate the population *at risk* of being issued a criminal summons, arrested, or having force used against them.[2] Only the presence of disparities can be calculated with benchmark analyses, not the presence of bias.

Finally, multivariate regression modeling is a statistical technique that creates a mathematical equation to estimate the influence of multiple variables on an outcome. While it is often convenient to focus on any given single factor that may affect officer decision-making, (e.g., subject's race) multivariate regression analyses are typically

[1] The seminal events examined using interrupted time series analyses include: (1) death of Elijah McClain, (2) Colorado Executive Order declaring COVID-19 Disaster Emergency, (3) officer-involved death of George Floyd in Minneapolis, (4) enactment of Colorado SB 20-217: Enhance Law Enforcement Integrity, (5) AG launch of pattern or practice investigation, (6) Independent Review Panel report released, (7) indictment of officers involved in McClain death, (8) city enters into Consent Decree, (9) APD Chief Vanessa Wilson terminated, and (10) Interim APD Chief Dan Oates hired.

[2] The NPI team used eight benchmarks in various analyses, including: (1) residential population data, (2) individuals issued criminal summonses, (3) all arrested individuals, (4) individuals arrested for Part I offenses, (5) individuals arrested for Part I violent offenses, (6) all crime suspects as reported to the police, (7) crime suspects for Part I offenses, and (8) crime suspects for Part I violent offenses.

considered more scientifically valid because these approaches quantify the impact of multiple factors simultaneously and estimate how confident we can be that the associations revealed are not due to random chance. Using multivariate regression analyses, the NPI team examined the likelihood of police use of force against subjects who have been arrested.

Combining statistical approaches allows for more comprehensive policy recommendations by understanding patterns and trends over time (descriptive analyses, interrupted time series analyses), addressing observed disparities (benchmark analyses), and identifying possible contextual factors that contribute to police enforcement actions (multivariate regression analysis). Although benchmarking is valuable for identifying and quantifying racial disparities, multivariate regression analyses supports the examination of the complex interplay of contributing factors. A holistic approach incorporating all statistical methods can offer a more comprehensive understanding of racial disparities in policing outcomes and inform effective policy interventions.

FINDINGS

It is important to note two caveats to the findings presented in this report. First, no statistical analysis using these data can determine if APD officers engage in *racially biased* enforcement actions. While it is possible to estimate *racial disparities* in enforcement actions (i.e., differences in outcomes across racial/ethnic groups) using a combination of statistical analyses, it is beyond the scope of any quantitative analysis to determine if the disparities observed are due to officer bias or discrimination.

Second, no *single* analysis can determine definitively if APD enforcement actions are racially disparate. Each type of statistical analysis has strengths and limitations that should be considered when interpreting the findings. It is possible that analyses employing different techniques or data sources produce conflicting findings. The purpose of conducting multiple analyses using a variety of data sources is to develop a more comprehensive understanding of APD enforcement patterns. With these caveats in mind, several notable findings are summarized below.

1) **Crime, especially serious and violent crime, steadily increased in Aurora from 2017 to 2022. This increase was not disrupted or accelerated by any seminal event examined.**

The increase in crime in Aurora across this six-year period includes a 20% increase in total criminal incidents, a 44% increase in reported criminal incidents involving Part I offenses, and an 82% increase in criminal incidents involving Part I violent offenses.

Time series analyses indicate a consistent upward trend of reported crime that was not significantly reduced or accelerated by seminal events, including the onset of the COVID-19 pandemic.

2) **As crime continued to increase from 2017 to 2022, the number of criminal summonses and arrests significantly decreased. This decline in enforcement activity was accelerated by the onset of the COVID-19 pandemic.**

The onset of the COVID-19 pandemic disrupted APD enforcement activities, significantly reducing the number of criminal summonses and arrests. No other seminal events were shown to significantly change criminal summonses and arrests. The significant increase in reported crimes in Aurora did not result in increased APD enforcement activity.

- APD officers issued 20,922 criminal summonses from 2017 to 2022, but there was a linear decline during this period. The onset of the COVID-19 pandemic significantly reduced the issuance of criminal summonses by an additional 11.2% (over and above the pre-established linear decline).
- APD officers arrested 44,954 individuals from 2017 to 2022. The onset of the COVID-19 pandemic significantly disrupted APD arrest activity, reducing it by approximately 50%. Decreases in arrests for less serious offenses were the primary drivers of this reduction. Therefore, although overall arrest counts decreased, the proportion of those arrests post-COVID for more serious and violent offenses increased.

3) **The annual number of subjects who had force used against them by APD officers remained relatively stable across the six-year period. This pattern was not interrupted by the COVID-19 pandemic, or any other seminal event examined.**

The number of individuals who had force used against them during this six-year period (total n = 3,783) did not significantly fluctuate annually.

- Despite stability in the number of individuals subjected to police use of force, the percentage of arrestees who had force used against them significantly increased. This was due to the decline in the number of arrests while use of force counts remained constant.

  Unlike both criminal summonses and arrests, use of force was not significantly disrupted by the onset of the COVID-19 pandemic. Data limitations prevented an in-depth analysis of factors contributing to the stability in use of force over time.

4) Substantively small to marginal racial/ethnic disparities in arrests were found using non-census benchmark comparisons.[3] For the most recent period (post-COVID), racial/ethnic disparities decreased and some benchmarks showed that Black and Hispanic individuals were *less likely* to be arrested compared to White individuals.

Of the 44,954 individuals arrested by APD officers, 40.3% were Black, 30.8% were White, 25.8% were Hispanic, and 3.1% were of other or unknown racial/ethnic backgrounds.

The NPI team compared the representation (%) of each racial/ethnic group in the population of arrested individuals to their representation in four different benchmarks, including (a) residential population, (b) all crime suspects, (c) crime suspects for Part I offenses, and (d) crime suspects for Part I violent offenses.

- Disparities in arrests for Black individuals compared to White individuals decreased post-COVID across all suspect-based benchmarks. For Hispanic individuals, two of the three suspect-based benchmarks also declined post-COVID. The post-COVID arrest disparity ratios based on Part I violent suspects show that both Black and Hispanic individuals were *less likely* to be arrested than their White counterparts.
- Although disparities in arrests for both Black and Hispanic individuals increased post-COVID when using the residential population-based benchmark, the validity of this benchmark (as an accurate measure of the population at risk of arrest) has been widely questioned and questioned by many experts.

5) Substantively small or no racial/ethnic disparities in uses of force were found using non-census benchmark comparisons. These small disparities were further reduced in the most recent time period (post-COVID).

Of the 3,783 individuals who had force used against them, 43.1% were Black, 33.5% were White, 15.3% were Hispanic, 5.7% were of unknown race/ethnicity, and 2.5% were other racial/ethnic backgrounds.

The NPI team compared the representation (%) of each racial/ethnic group in the population of those who experienced force to their representation in eight different benchmarks, including (a) residential population, (b) criminal summonses, (c) all

[3] All benchmarks have limitations and vary in the extent to which they accurately estimate the population of individuals "at risk" of police enforcement actions. Based on research regarding the validity of different benchmarks and the factors that influence police behavior, criminal suspect-based benchmarks are considered stronger approximations of the population at risk of arrests or use of force compared to other benchmarks, while residential census data is widely considered an unreliable and invalid comparison measure (Alpert et al., 2004; Fridell, 2004; Geller et al., 2021; Smith et al., 2019) and arrest data may mask or underestimate racial/ethnic disparities (Knox et al., 2020a, 2020b).

arrestees, (d) arrestees for Part I offenses, (e) arrestees for Part I violent offenses, (f) all crime suspects, (g) crime suspects for Part I offenses, and (h) crime suspects for Part I violent offenses.

- Benchmark analyses for use of force show small or no racial/ethnic disparities for Black individuals in use of force across most of the eight benchmarks examined.
- Disparities in use of force for Black individuals compared to White individuals decreased post-COVID across all benchmarks, while no disparities in use of force for Hispanic individuals were evident across the benchmarks either before or after the onset of COVID.
- As with arrests, only the residential population benchmark demonstrated racial/ethnic disparities in police use of force, and only for Black individuals compared to White individuals.

6) **When examining only arrestees, multivariate analyses show that Black arrestees are significantly more likely to have force used against them compared to White arrestees after controlling for other situational, legal, and arrestee characteristics. Hispanic arrestees are *not* significantly more likely to experience force compared to White arrestees.**

Although the differences in the likelihood of use of force for Black compared to White arrestees is statistically significant, it represents a substantively small difference in the predicted probability of use of force. These racial differences are also reduced in the post-COVID period compared to the approximately three years prior.

- The multivariate analyses also show that Hispanic arrestees were *not* significantly more likely to experience a use of force than White arrestees during the six-year period after controlling for other situational, legal, and arrestee characteristics.
- The results of the multivariate analyses must be interpreted cautiously because the strongest known predictors of use of force (e.g., suspect resistance, intoxication, presence of a weapon, etc.) could not be included in the statistical models due to limitations in the available arrest data.

7) **Collectively, the analyses suggest that any differences in APD enforcement actions across racial/ethnic groups are small to marginal, and disparities that initially exist have significantly declined over time.**

Taken as a whole, the statistical analyses examining racial/ethnic disparities in APD enforcement actions are small in magnitude, reducing over time, and in some cases, do not exist in the data analyzed. Note however that these analyses have methodological and data quality limitations and should, therefore, be interpreted with caution. The best use of this information is to establish a series of repeated measures to explore the impact of police reform efforts over time.

RECOMMENDATIONS

Based on the findings reported above, the NPI team provides the following five recommendations for APD's continued improvement in data collection, policy, training, and operational enforcement practices.

Recommendation 1: Continue data collection system overhaul.

The APD has been actively developing a new system for reporting and collecting use of force data that should be operational soon. Improvements to the reporting system will assist in better understanding the dynamics of use of force interactions, exploring whether there are racial/ethnic differences in correlates of use of force, and examining the factors that predict subject and officer injuries, all of which can potentially inform additional improvements to use of force policy and training.

Recommendation 2: Add more accountability checks for accurate data collection to demonstrate its importance.

For APD to continue to be data-driven in its practices and to provide transparency to the community, the department must improve the quality of its use of force data. The APD should develop or enhance reliability and validity checks, including validation measures within the data reporting system, APD's chain of command review processes, and periodic data audits.

Recommendation 3: Continue updates in UOF policy and training.

As part of its ongoing effort to update policies and procedures, the APD should consider revising Directive 05.05 *Reporting Use of Force* to reclassify the pointing of a firearm from Tier Zero to Tier One. This would facilitate more detailed reporting and evaluation by supervisors and commanders to ensure these actions align with department policy and reduce the risk of accidental or unjustified shootings.

In 2023, the APD trained its personnel in the Police Executive Research Forum's (PERF) Integrating, Communications, Assessment, and Tactics (ICAT) de-escalation training. The NPI team recommends that the APD implement strategies for maximizing and sustaining the benefits of de-escalation training, as outlined in a recent PERF implementation guide.

Recommendation 4: Continue to track changes in racial/ethnic disparities in APD enforcement actions using multiple measures and analytical techniques.

Determining whether racial/ethnic disparities exist in enforcement actions can be complex but is necessary for guiding any law enforcement agency's approach to addressing them. The information provided in this report should be used by the Independent Consent Decree Monitor, the City, and the APD to assist the department in meeting consent decree mandates and aligning with best practices. The APD should consider partnering with an independent research team to continue this work.

The findings presented within this report are based on multiple data sources and statistical techniques. Rather than estimating the amount of racial/ethnic disparity in APD enforcement activities, these findings are better used as baseline measures for comparisons over time. Regardless of the specific level of disparities – which vary based on the data used and analyses conducted – progress toward the reduction of disparities over time can be estimated as reforms are implemented.

Recommendation 5: Implement effective and equitable crime reduction strategies immediately – especially focused on violence – and continually monitor the impact on reported crime, enforcement disparities, and community sentiment.

It is critical for the APD and the City of Aurora to implement strategies that can effectively address the rise in violent crime without exacerbating racial disparities in APD enforcement outcomes or sacrificing community trust in the police. Specific consideration should be given to evidence-informed, place-based, and individual-oriented strategies to address factors that contribute to violent crime. Implementing a comprehensive, city-wide violence prevention strategy focusing on the highest-risk places and people can help Aurora reduce violence while maintaining positive reductions in racial disparities across policing outcomes.

CONCLUSION

This report provides baseline measures for examining racial disparities in enforcement against which the APD can compare future years of data. However, the findings should be interpreted with caution. Regardless of the available data or statistical analyses employed, this aggregate, quantitative examination of patterns and trends in enforcement outcomes cannot determine whether APD officers have made enforcement decisions based on racial bias. Data collection and analyses, however, can provide police executives with the necessary information to examine potentially problematic areas more closely and identify opportunities for improvement where warranted. It also demonstrates transparency to the public and commitment toward evidence-based policing practices that can help to make police encounters with the public safer and more equitable.

SECTION 1: INTRODUCTION

On November 16, 2021, the Colorado Office of the Attorney General announced its initiation of a consent decree with the City of Aurora, Colorado following the recommendations of an investigative team appointed by Colorado's Attorney General, Phil Weiser, to conduct a pattern or practice investigation. The investigative report, released on September 15, 2021, documented the APD's engagement in activities related to racially biased policing, the use of excessive force, and the failure to record pertinent information in officers' interactions with community members (see Weiser, 2021, p. 1, par. 1). The consent decree mandated oversight of the APD, as well as the Aurora Fire Rescue and Aurora Civil Service Commission, with all three agencies ordered to amend current policies, procedures, and training to increase public trust, enhance the legitimacy and transparency of emergency services, and advance community safety in Aurora.

On February 14, 2022, IntegrAssure, LLC., was appointed as the Independent Consent Decree Monitor to oversee these agencies' implementation of consent decree mandates and ensure progression toward compliance goals that align with state and federal laws. To support their monitorship, IntegrAssure engaged the National Policing Institute (NPI) to conduct statistical analyses and interpret APD enforcement data to develop baseline measures that may be used to examine changes in police activity and outcomes as the APD implements efforts to meet the consent decree requirements. This work will facilitate IntegrAssure's assessment of whether the City has changed "in measurable ways, how Aurora Police engages with all members of the community, including by reducing any racial disparities in how Aurora Police engages, arrests, and uses forces in the community" (Consent Decree, 2022, p.7).

Current Work

In April 2023, the NPI team produced a technical report describing the research plan to establish baseline measures for the City, including a description of the data sources, methodologies, and statistical techniques to be used. The current report presents the findings from the NPI team's examination of the patterns and trends in the APD's enforcement activities over time (2017-2022). Using multiple data sources and analytic approaches, this report outlines key baseline measures across reported criminal offenses, criminal summonses, arrests, and uses of force that may be used for comparison in future examinations of the APD's enforcement practices and racial/ethnic disparities in enforcement. This report is organized as follows:

Section 2 identifies the research questions, data sources, and analytic strategies used to examine APD's enforcement activities over time. This description includes the strengths and limitations of the data, measures, and analytic strategies used by the NPI team. **Section 3** examines trends and patterns in reported criminal offenses, criminal summonses, arrests, and use of force over time. **Section 4** presents benchmark comparisons of the rates of arrest and use of force experienced by different racial/ethnic groups to different comparison populations. **Section 5** presents findings from analyses examining the predictors of APD officers' use of force during arrests. Finally, **Section 6**, summarizes the main findings of the report and provides recommendations for the Independent Consent Decree Monitor and the APD to consider opportunities to continuously improve use of force data collection, policy and training and to promote community and officer safety.

SECTION 2: METHODOLOGY

This section presents the research questions guiding the NPI team's data collection and analytic strategy and describes the data sources used to examine APD's enforcement activities. An overview of the main techniques used in the analysis plan is also provided, along with a comprehensive assessment of the reliability and validity of the available data.

RESEARCH QUESTIONS

The data collection and analytic strategy for this work were guided by six research questions pertaining to the enforcement activities of the APD over time and across groups of community members. These questions included:

(1) What are the trends and patterns in APD's criminal summonses, arrests, and uses of force over time?

(2) Does the frequency of criminal summonses, arrests, and uses of force shift significantly after seminal events (i.e., events at discrete points in time believed to influence police-citizen encounters)?

(3) Do rates of arrest and use of force experienced by different racial/ethnic groups align with their representation among the populations at risk of experiencing enforcement actions by the APD?

(4) What factors predict the likelihood of use of force by APD officers?

(5) Are community members' race, ethnicity, or gender related to the type or severity of the force used by the police?

(6) What factors predict the likelihood of injuries to community members or officers during use of force incidents?

Notably, research questions 5 and 6 could not be answered due to limitations in the data collected by the APD. It is unknown whether individuals' demographic characteristics are related to the type or severity of force used by APD officers. Additionally, the factors contributing to the likelihood of injury to community members or police officers during use of force incidents cannot be identified. These limitations are discussed further below.

DATA DESCRIPTION

Several official APD data sources were used to triangulate findings and provide a holistic understanding of the factors influencing police enforcement actions. The primary APD data sources include:

(1) Reported criminal offenses (including suspect information, when available)
(2) Criminal summonses
(3) Arrests
(4) Uses of force

Electronic data was received from the City of Aurora and the APD for six consecutive years: January 1, 2017–December 31, 2022.

Criminal Offense Data

The reasons to consider criminal offense data in the development of baseline measures for APD's enforcement activities are two-fold. First, the review of reported criminal offenses allows for the examination of trends in criminal incidents over time, providing important context for APD's enforcement activities. Second, the examination of reported criminal offenses supports the identification of benchmark populations for known criminal suspects (described later in this section). These benchmark populations facilitate comparisons of rates of enforcement activities experienced across groups to understand if racial disparities exist.

Table 2.1 displays the measures used by the NPI team from the criminal offense data provided by the APD. These data identify 49,173 criminal offenses reported to the APD during the six-year period of interest. Criminal offense data contain information about the reported criminal incident and the suspect (if known). A single criminal incident can involve more than one offense. Similarly, a single incident can involve more than one suspect. For the present analyses, reported criminal offense data were aggregated to the incident and suspect levels using the incident number, date, and suspect identifier. This aggregation identifies 35,889 individuals involved in 33,495 criminal incidents over the six-year period.

Offense data fields were used to create crime-type categories. The variables created for Part I crimes and Part I violent crimes should be interpreted as the percentage of incidents or suspects with at least one Part I offense or Part I violent offense.

Table 2.1. Available Measures in APD Criminal Incident Data

Variable Name	Description	Recoded Variables Used in Analyses
Month of Criminal Incident[4]	Jan–Dec Incident Dates (by Month)	1 = Jan, 2 = Feb, 3 = Mar, 4 = Apr, 5 = May, 6 = Jun, 7 = Jul, 8 = Aug, 9 = Sep, 10 = Oct, 11 = Nov, 12 = Dec Q1 = Jan-Mar, Q2 = Apr-Jun, Q3 = Jul-Sep, Q4 = Oct-Dec
Year of Criminal Incident	Six years: 2017–2022	Numeric value of year
Time of Day	Time of incident collected using the 24-hour clock	Binary variable *Daytime incident* 0 = night (7:00 PM–6:59 AM) 1 = day (7:00 AM–6:59 PM)
Location	Latitude and longitude, or street address of criminal incident	Latitude and longitude or street address used to geocode/map crime incidents
Criminal Offense	Charge/offense and associated UCR codes	348 charge codes, recorded into two categories: Part 1 violent = aggravated assault, rape, robbery, murder Part 1 overall = larceny, burglary, motor vehicle theft, and all Part 1 violent
Suspect Gender	Gender of the criminal suspect	Binary variable, 0 = female, 1 = male
Suspect Race/Ethnicity	Race/ethnicity of criminal suspect; original race categories include: White, Black, American Indian or Alaskan Native, Asian or Pacific Islander, unknown Original ethnicity categories include: Hispanic, not Hispanic, unknown	*Race/ethnicity* 1 = White (non-Hispanic) 2 = Black (including Hispanic Black) 3 = Hispanic (including White or unknown race) 4 = Other (American Indian or Alaskan Native, Asian or Pacific Islander, unknown)
Suspect Age	Age of suspect at time of incident	Continuous variable measured in years between date of birth and criminal incident date

Criminal Summons Data

APD uses both physical (or custodial) arrests, where a person is taken into police custody, and criminal summonses, where a person is issued a summons to appear in court but is not taken into custody. The criminal summons data are described below, followed by the custodial arrest data.

For the six-year study period (2017–2022), the summons data provided to the NPI team included 91,990 rows of data corresponding to each charge rather than each person

[4] The incident date and time fields were received in a short text data format. They were converted to a date and time format to facilitate time series analyses by creating monthly counts.

charged.[5] This information was aggregated to the individual (person-charged) level. Once aggregated, 57,586 individuals[6] received summonses during the six-year period; however, 64% (36,664 individuals) were excluded from analyses because they were issued traffic summonses with no criminal charge.[7] In total, the criminal summons data examined in Section 3 is based on 20,922 individuals issued criminal summonses from 2017–2022 by APD officers.

These data allow for the examination of the trends in the number of individuals issued criminal summonses over time and are used as a benchmark population comparison for those who experienced force. **Table 2.2** displays the variables and recoded measures for the criminal summons data, including incident characteristics, legal factors, and demographic characteristics of individuals issued summonses.

[5] That is, a person issued a summons for a single charge had one row of data, whereas a person issued a summons for three charges had three rows of data.
[6] Individuals may appear in the dataset more than once for different incidents.
[7] Those issued only traffic summonses were excluded from this report because there are no analyses included for traffic stops (data not collected by APD during the study period).

Table 2.2. Available Measures in APD Criminal Summons Data

Variable Name	Description	Recoded Variables Used in Analyses
Month of Summons	Jan–Dec Incident Dates (by Month)	1=Jan, 2=Feb, 3=Mar, 4=Apr, 5=May, 6=Jun, 7=Jul, 8=Aug, 9=Sep, 10=Oct, 11=Nov, 12=Dec Q1=Jan-Mar, Q2=Apr-Jun, Q3=Jul-Sep, Q4=Oct-Dec
Year of Summons	Six years: 2017–2022	Numeric value of year
Day of the Week	Day of week summons issued	Binary variable *Weekend* 0 = work week (Mon-Thu) 1 = weekend (Fri-Sun)
Time of Day	Time of summons, collected using 24-hour clock	Binary variable *Daytime incident* 0 = night (7:00 PM–6:59 AM) 1 = day (7:00 AM–6:59 PM)
Multiple Subjects	Incident involved more than one person issued criminal summons	Binary variable *Multiple_People* 0 = single person issued summons 1 = multiple people in single incident issued summonses
APD District	APD patrol district where summons issued, based on incident address/location	*APD district* 0 = missing or out of district, 1 = District 1, 2 = District 2, 3 = District 3
Subject Gender	Gender of the criminal suspect	Binary variable, 0 = female, 1 = male
Subject Race/Ethnicity	Race/ethnicity of the person issued a summons. Original race categories include: White, Black, American Indian or Alaskan Native, Asian or Pacific Islander, unknown Original ethnicity categories include: Hispanic, not Hispanic, unknown	*Race/ethnicity* 1 = White (non-Hispanic) 2 = Black (including Hispanic Black) 3 = Hispanic (including White or unknown race) 4 = Other (American Indian or Alaskan Native, Asian or Pacific Islander, unknown)
Subject Age	Age of subject at time of summons	Continuous variable measured in years between date of birth and criminal summons date

Arrest Data

APD policy requires that custodial arrests result in documentation (i.e., arrest reports), including arrestee demographic characteristics, specific criminal charges, and some situational characteristics of the incident. These arrest data are collected and stored in the APD's Versadex data management system. For the study period (2017–2022), the data provided to the NPI team included 46,932 rows of data, corresponding to each specific criminal charge for all arrestees.[8] This information was aggregated to the

[8] For example, an individual arrested for a single charge had one row of data, whereas an individual arrested for three charges had three rows of data.

individual arrestee level for a total of 44,954 individuals arrested.[9] The majority of APD arrests involve individuals with a single charge, as only 1,753 (3.9%) of those arrested have two or more charges per custodial event.

The arrest data are used for four purposes: (1) to examine arrest trends over time and across arrestees' race/ethnicity, (2) to examine racial/ethnic disparities in arrests with benchmark analyses, (3) to create benchmark populations for comparisons to those who experienced force, and (4) to understand the factors that predict whether arrestees have force used against them. For the last analysis, arrestees and uses of force are linked by a unique identifier, where applicable.

Table 2.3 displays the variables and recoded measures for the arrest data, including incident characteristics, legal factors, and arrestee demographics. Notably, the arrest data do not include relevant additional information, including arrestees' compliance or resistance, mental health considerations, drug or alcohol use, or presence of a weapon (absent a weapon charge). These situational factors (unmeasured in the APD data) have routinely been shown as the strongest predictors of officer use of force.[10]

[9] A single person could be arrested multiple times over the six-year period, and in those cases, each arrest for the person counts as an individual-arrest (multiple dates for the same person). Thus, the total of 44,954 individual arrestees is not equivalent to 44,954 *unique* individuals arrested. Additionally, multiple individuals could be arrested in a single incident (on the same date/time), and in those cases, each person-incident is also counted separately.

[10] For example, see Engel, 2015; Engel et al., 2000; Engel & Swartz, 2014; Garner et al., 2002; Gau et al., 2010; Kramer & Remster, 2018; Rossler & Terrill, 2017; Smith et al., 2022; Stroshine & Brandl, 2019; Terrill & Mastrofski, 2002.

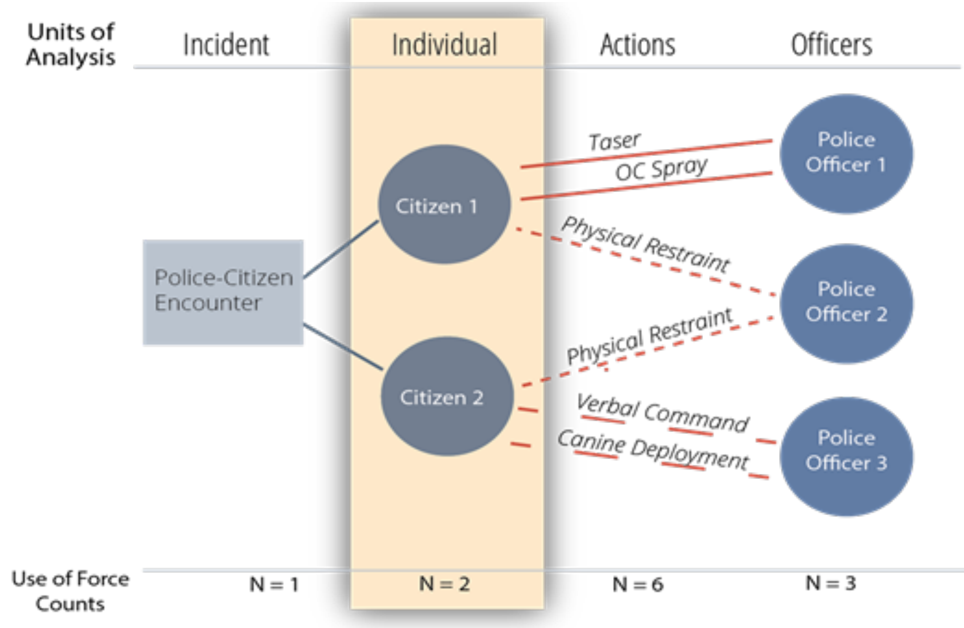
Table 2.3. Available Measures in APD Arrest Data

Variable Name	Description	Recoded Variables Used in Analyses
Month of Arrest	Jan–Dec Incident Dates (by Month)	1=Jan, 2=Feb, 3=Mar, 4=Apr, 5=May, 6=Jun, 7=Jul, 8=Aug, 9=Sep, 10=Oct, 11=Nov, 12=Dec Q1 = Jan-Mar, Q2 = Apr-Jun, Q3 = Jul-Sep, Q4 = Oct-Dec
Year of Arrest	Six years: 2017–2022	Numeric value of year
Day of the Week	Day of week arrested	Binary variable *Weekend* 0 = work week (Mon-Thu) 1 = weekend (Fri-Sun)
Time of Day	Time of arrest collected using 24-hour clock	Binary variable *Daytime incident* 0 = night (7:00 PM–6:59 AM) 1 = day (7:00 AM–6:59 PM)
Outstanding Warrant	Arrestee has outstanding warrant	Binary variable *Outstanding Warrant* 0 = No outstanding warrant 1 = Outstanding Warrant
Violent Offense	Charges against arrestee include Part I violent offense charges	*Violent Offense Arrest*: 0 = no violent offense charges 1 = At least one charge for Part I violent offense
Multiple Arrestee	Incident involved more than one person arrested	Binary variable *MultiArrest* 0 = single arrestee 1 = arrestee one of multiple arrestees
APD District	APD patrol district where arrest occurred, based on incident address/location	*APD district* 0 = missing or out of district, 1 = District 1, 2 = District 2, 3 = District 3
Arrestee Gender	Gender of arrestee	Binary variable, 0 = female, 1 = male
Arrestee Race/Ethnicity	Race/ethnicity of arrestee. Original race categories include: White, Black, American Indian or Alaskan Native, Asian or Pacific Islander, unknown Original ethnicity categories include: Hispanic, not Hispanic, unknown	*Race/ethnicity* 1 = White (non-Hispanic) 2 = Black (including Hispanic Black) 3 = Hispanic (including White or unknown race) 4 = Other (American Indian or Alaskan Native, Asian or Pacific Islander, unknown)
Arrestee Age	Age of subject at time of arrest	Continuous variable measured in years between date of birth and arrest date

Use of Force Data

When APD officers use force against individuals, their supervisor is required to complete a Use of Force Report. These reports are completed in APD's AIM (Administrative Investigations Management) system and include information about the incident, the involved officer(s), the subject(s), the reason for force, force actions used, and any resulting injuries to the officers or subjects. For analytical purposes, information from the force reports must be aggregated to the unit of analysis of interest. The possible units of analysis are graphically displayed in **Figure 2.1**.

Figure 2.1: Hypothetical Example of Use of Force Measures by Unit of Analysis



Note: The individual level is the unit of analysis used in this study.

Source: Engel, R. S., Corsaro, N., Isaza, G. T., & McManus, H. D. (2022). Assessing the impact of de-escalation training on police behavior: Reducing police use of force in the Louisville, KY Metro Police Department. *Criminology & Public Policy*, p.211.

The counts of use of force can vary dramatically depending on which unit of analysis is selected (Engel et al., 2022). For example, in the example above, there is one incident, two subjects or individuals, six force actions, and three officers. This distinction in units of analysis is noted in APD's publicly available use of force reports.[11] For all analyses in this report, use of force is analyzed at the *individual/subject level* (i.e., an individual who experiences a use of force within a single incident) given the interest in examining differences in uses of force across individuals of different racial/ethnic groups.

Analyses at the subject level required aggregating information from multiple use of force data tables at various levels within the APD's AIM database to the subject level for all use of force incidents. As described in detail below, this process of aggregating information from different units of analysis to the subject level was analytically

[11] For example, see the 2020 Annual Use of Force Report that included 500 use of force incidents, 644 subjects, 413 officers, and 1,579 applications of force. https://cdnsm5-hosted.civiclive.com/UserFiles/Servers/Server_1881137/File/Residents/Public%20Safety/Police/Public%20Reports%20and%20Crime%20Data/2020%20Annual%20Use%20of%20Force%20Report.pdf

challenging. Ultimately, for the six-year study period, the use of force data includes 3,783 individuals involved in 3,518 incidents.

Table 2.4 displays the variables and recoded measures for the use of force data, including incident characteristics, type of force used, demographics of the individuals against whom force was used, and injuries. Notably, despite resistance being defined in the APD Use of Force Glossary, a measure of subject resistance was not captured in the data provided to the NPI team. The APD is actively revising the department's use of force data collection, with information on individuals' resistance to be systematically captured in the updated use of force reporting protocol.[12]

[12] Current definitions of subject resistance can be found in APD Directive 05.01 *Use of Force* at https://public.powerdms.com/AURORAPD/tree/documents/3167288

Table 2.4. Available Measures in APD Use of Force Data

Variable Name	Description	Recoded Variables Used in Analyses
Incident ID	Unique numeric identifier to link use of force files	Measured as collected (string identifier)
Day of the Week	Day of week arrested	Binary variable *Weekend,* 0 = work week (Mon-Thu), 1 = weekend (Fri-Sun)
Time of Day	Time of incident collected using 24-hour clock	Binary variable *Daytime incident,* 0 = night (7:00 PM–6:59 AM), 1 = day (7:00 AM–6:59 PM)
Address of incident	Incident address linked to the various geographic files	Each location has a unique identifier
Subject gender	Gender of use of force subject	Binary variable, 0 = female, 1 = male
Subject race/ethnicity	Use of force subject race/ethnicity Original race/ethnicity categories include: White, Black, Hispanic[13], American Indian, Asian, Hawaiian or Pacific Islander, Mixed, Other, Unknown	Subject Race/ethnicity coded as 1 = White, 2 = Black, 3 = Hispanic and 4 = Other (including all other categories)
Subject age	Age of subject at time of force	Continuous variable measured in years between date of birth and use of force date
Subject alcohol or drug impairment	Use of force subject perceived to be impaired by alcohol or drugs	Binary variable *Substance Impairment,* 0 = not impaired, 1 = impaired by drugs or alcohol
Reason for force	Officer's legal justification for using force	Necessary to: 1= effect arrest, 2 = prevent a crime 3 = defend another, 4 = defend officer 5 = prevent a crime, 6 = for subjects' safety, 7 = failure to obey
Type of force	Officer-level input, aggregated to the individual-incident level (i.e., if any officer used any of the following actions in the incident against the person)	1 = control techniques (twist locks, takedowns, etc.), 2 = hobble, 3 = O.C. spray, 4 = punches, strikes, kicks, 5 = Baton, 6 = Taser, 7 = police canine, 8 = launchable impact weapons, 9 = Other (i.e., PIT maneuver, stop sticks), 10 = Deadly force
Type of offenses	Offense use of force subject charged with	1 = Misdemeanor, 2 = Felony, 3 = Protective custody, 4 = Petty offense
Subject arrested	Use of force subject arrested	Binary variable coded as 0 = No, 1 = Yes
Subject injury	Use of force subject injured	Binary variable coded as 0 = No, 1 = Yes
Subject preexisting injury	Use of force subject preexisting injury	Binary variable coded as 0 = No, 1 = Yes
Subject treatment	Use of force subject provided medical treatment	0 = not needed; 1 = Treated/release; 2 = Professional medical treatment; 3 = Hospitalized
Officer injured	Officer injured in use of force incident	0 = No injury; 1 = Injury
Officer (injury) treatment	Type of treatment officer injury	0 = Treatment not needed; 1 = Treated/released; 2 = Hospitalized

[13] Unlike arrest data where race and ethnicity are two separate data fields, subject race/ethnicity is captured as a single race/ethnicity field in the use of force report.

Table 2.5 displays the tiered system the APD uses to classify types of force. The reporting, investigatory, and review processes vary by force tier. Most types of force were introduced in APD policy as "Incidents that Require Notification and Reporting" on January 3, 2015, without the associated tiers.[14] On January 1, 2016, Tier Zero force types were introduced in APD policy, and the preexisting types of force were categorized into the tiers the APD presently uses. All other types of force introduced later are noted with their effective date in parentheses in Table 2.5. The APD released an updated version of its Reporting Use of Force policy on August 18, 2023.[15]

This report only examines Tiers One to Three uses of force; Tier Zero incidents do not result in a report. As a result, the frequency or patterns and trends associated with the APD's use of pointing a firearm cannot be assessed. By way of comparison, recent research examining use of force by the Colorado Springs Police Department found that the pointing of a firearm comprised approximately two-thirds of the department's use of force (Brown et al., 2022).

[14] APD Policy DM 05.04 - Reporting and Investigating the Use of Tools, Weapons, and Physical, p.1, Section 5.4.1, 2015.
[15] Table 2.5 differs from that shown in the *Technical Report* (issued in April 20223), as it has been updated to reflect policy changes.

Table 2.5. APD Types of Force by Tier

Tier Level	Types of Force Included
Tier Zero Statutory use of force per C.R.S. § 18-1-707 or display of force by a sworn member of APD	Firearm Gun Point (Handgun, Rifle, Shotgun)
	Less Lethal Shotgun, Projectile Launcher, Taser, OC Pointing
	Arrest with Handcuffs (Introduced as Tier Zero 10/07/2020)
	Handcuff and Release with no arrest or summons
	Pat-Down for weapon (Introduced as Tier Zero 10/07/2020)
	Physically redirecting a person that does not involve overcoming resistance (Introduced as Tier Zero 10/07/2020)
	Consensual Search of a Person (Introduced as Tier Zero 10/07/2020)
Tier One Use of force with no or minor injury used to overcome physical resistance	Take Down (Introduced as Tier One 01/01/2016)
	Use of control weapon (Baton or SD-1) for leverage or control purposes (no strikes or thrusts)
	BolaWrap™ (Introduced as Tier One 08/18/2023)
	WRAP™ or Restraint Chair
	Restraining measures to assist AFR, EMS, and/or medical personnel
Tier Two[16] Use of a weapon other than a deadly weapon or actions that result in injury requiring professional medical treatment	Oleoresin Capsicum (pepper spray)
	Baton Strikes/Thrusts
	Launchable Impact Weapons
	CEW
	Use of Personal Weapons (e.g., strikes, punches, kicks)
	Police Canine Sent with the Intent to Bite
	Pitting and/or Boxing of a Moving Vehicle (Introduced as Tier Two 01/01/2016)
	Tire-deflation device used on a vehicle in motion with successful tire deflation (Introduced as Tier Two 10/07/2020)
Tier Three[17] Use of a deadly weapon, lethal force, and/or force where hospitalization or death occurs	Use of Lethal Force regardless of injury[18]
	Use of force, tools, or weapons which result in hospitalization or death[19]
	Intentional use of a vehicle against a person on foot (Introduced as Tier Three 08/18/2023)
	Any incident where a sworn member discharges a firearm and a person is struck by a bullet outside of a training environment

Source: Adapted from APD's UOF Matrix (Vers 4), APD's DM 05.04 – Reporting and Investigating the Use of Tools, Weapons, and Physical (Vers 0-14), and APD's 05.05 Reporting Use of Force Policy

[16] Carotid Control was classified as a Tier Two type of force on Jan 1, 2016, and was **prohibited** from use on Jun 9, 2020. Therefore, it is excluded from the table that reflects current tiers of force.
[17] When a supervisor, in conjunction with the Duty Executive, believes that a use of force warrants a higher or lower Tier Classification and response, they can adjust accordingly.
[18] Effective Dec 7, 2016, "except in incidents involving a firearm, when the use of a tool or weapon that is considered potentially deadly force is used to overcome resistance resulting in no injury, or injury not requiring professional medical treatment, the Duty Captain, in consultation with the Duty Executive, may direct that the incident be investigated as a Tier Two use of force."
[19] Effective May 13, 2019, "when a person is hospitalized due to use of force that would otherwise be considered a Tier Two use of force such as but not limited to Taser, K9 or less lethal deployment, the Duty Executive may determine that a Tier Two response (including all reporting) is appropriate."

DATA LIMITATIONS

Significant data limitations and quality concerns constrained the NPI team's capacity to compile and analyze the data provided by the APD to support this study. The NPI team collaborated directly with the APD and City of Aurora IT personnel to collect and prepare the required data. The data collection and transfer process involved multiple requests, various iterations of data submission, and a series of conversations with APD and City employees to extract necessary information and document how datasets are maintained and matched across the APD's systems. Once data was received, the preparation process involved merging multiple data tables and manually cleaning numerous free text fields.

The comprehensive data collection and management process for data relating to reported criminal offenses, criminal summonses, and arrests provided the NPI team with an acceptable level of confidence in the reliability of the data and subsequent analyses of the outcomes of interest. In contrast, the NPI team identified fundamental issues with the APD's use of force data that prohibited many originally planned analyses, detailed below. Notably, the NPI team cannot speak to the quality of the use of force data stored in the APD's original PDF Use of Force Reports. The examination detailed below is based only on the electronic data that was extracted by APD personnel for the purpose of this study.[20]

Linking Officers and Subjects

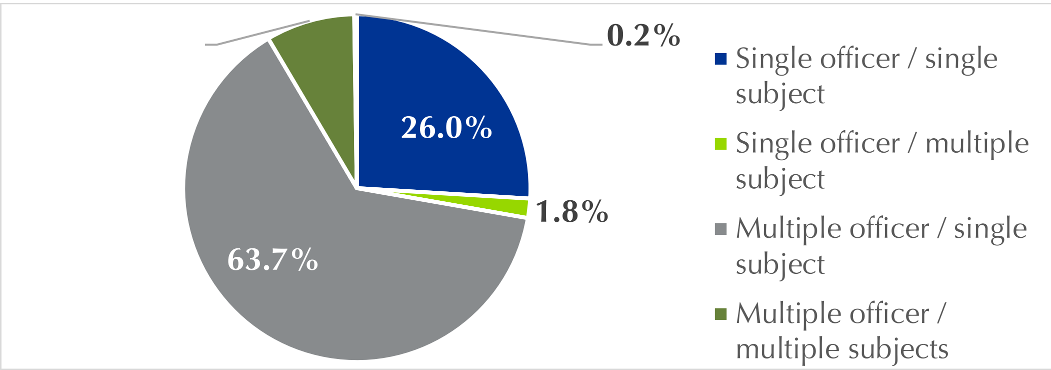
Using the APD's electronically available use of force data, the NPI team was unable to consistently link use of force subjects to the officer who used force against them. Use of force information was provided to the NPI team in multiple tables that had to be manually linked by incident number and individuals' unique identifiers. Notably, the APD's AIM system includes an "employee person link" that permits an analyst to connect the specific force actions and resulting injuries between each officer and subject for each force event. Unfortunately, this link is not reliably available for all cases. More than a third (n = 1,291, 34.1%) of the 3,783 individuals who had force used against them could not be reliably linked to the officers who used force. An examination of the data by year indicates that the issue with missing linkage information improved significantly over

[20] The only exception includes the NPI team sending a list to the APD of 23 incident numbers to assess whether the data issues identified were related to the original data reported or the data extraction process. Using the original AIM reports provided to the monitor, the NPI team compared the provided data with the original reports for this small sample. The findings of these comparisons are incorporated into the discussion below of each data issue discovered throughout the data preparation process.

time, declining from one-third of the cases in 2017 that could not be linked to less than 1% of cases in 2022.[21]

The prevalence of multi-officer use of force incidents further complicated the capacity to link officers and subjects. As displayed in **Figure 2.2**, only 26% of individuals (n = 982) who experienced force were involved in a single officer, single subject event. The remaining majority (over 70%, n = 2,754) were involved in a multi-officer event.

Figure 2.2. Distribution of Officers and Subjects Within Use of Force (n = 3,783)



The problem with linking officer and subject information in these incidents is best illustrated by the data fields related to "reason for force" and "type of force." Reason for force is missing for only 53 of the 3,783 individuals who had force used against them (1.4%). However, it is captured at the incident level rather than the officer or subject levels. As such, the NPI team could only document the reason for force recorded by all officers against all subjects within the incident rather than for each individual officer's application of force against individual subjects. While there is likely to be a high level of uniformity in the reason for force across single-subject, multi-officer incidents, this may not necessarily be the case for the roughly 10% of individuals involved in multiple-subject use of force incidents. Based on the available linkage information, the NPI team could not identify the reason for force for 39.4% of the 3,783 individuals who had force used against them.

Type of force information is collected at the subject-officer level. The "employee person link" described above connects each officer to each subject and lists the type of force used. Unfortunately, this information was only available for 65.9% of the cases. Furthermore, even among the linked cases, the type of force was missing for 160 individuals. Therefore, only 61.6% of the individuals who experienced force had reliable information to identify the specific type of force used against them. Finally, type of force was also included in the incident level data provided to the NPI team, but this

[21] Of the 1,291 cases where officer and subjects could not be linked, 33.0% occurred in 2017, 25.7% in 2018, 21.9% in 2019, 13.6% in 2020, 5.0% in 2021, and 0.8% in 2022.

information aggregated all types of force across officers and subjects, so it cannot be presumed to be accurate for individuals involved in multiple-subject incidents.

Missing Data

The use of force data also had many missing values across multiple measures necessary for substantive analyses of use of force. The missing information for various data fields is displayed in **Table 2.6**. Missing data for many of these measures is attributable to original reports not being fully completed and problems with the data extraction.

Table 2.6. Missing Data in Use of Force Data (n = 3,783)

Fields	% Missing
Date	0.0%
Time	5.9%
Location	27.8%
Subject Date of Birth (Age)	5.8%
Subject Race	5.5%
Subject Gender	4.8%
Subject Alcohol impairment	76.8%
Subject Drug impairment	76.6%
Subject Arrested	44.0%
Subject Injured (Yes/No)	75.6%

For example, when considering where the use of force incident occurred, 27.8% of the individuals had missing data for the location of their use of force incident (n = 1,051). In reviewing the sample of AIM reports, the NPI team discovered two separate fields for "location" and "address." Based on a review of the sampled cases, these fields are used interchangeably by APD personnel.[22] Unfortunately, the data pulled for the NPI team only included the "location" field. It is unknown why both fields are included in the report or how the APD personnel are trained to complete the "location" and "address" data fields, but there are inconsistencies in their use.

Based on the 2,732 use of force subjects with provided location data, the NPI team was able to geocode the incidents using street addresses for 1,840 (a 67.3% hit rate). The NPI team then geocoded an additional 1,301 cases (34.4% of the total dataset) based on address information included in the arrest and crime incident datasets after linking. Finally, an additional 216 cases were identified for manual geocoding via Google Maps and partial information located in the raw data. In total, 426 cases (11.3%) could not be

[22] For example, street addresses may be included under "location" with "address" missing and vice versa. In addition, sometimes a location type or business name is provided in the "location" field while the street address is reserved for the "address" field.

geocoded, equating to a total geocoding hit rate of 88.7%, which is above the typical minimum standard (Ratcliffe, 2004).

Similar missing data issues were discovered for subjects' characteristics. As shown in **Table 2.6,** the missing data for subjects' age, race, and gender was approximately 5-6%. This may be due to the use of force within crowd control situations. The remaining subject data fields are considerably more problematic. More than 75% of the individuals who had force used against them had missing data regarding whether the officer perceived them to be impaired by drugs and/or alcohol.[23]

The data field indicating whether a subject was arrested during a use of force incident is unreliable as it is missing for 44% of the individuals. As a result, the NPI team manually linked the use of force subjects with arrest data by using the incident number and unique subject identifier. Not everyone with force used against them was arrested, but the missing data cannot be presumed to be equivalent to "no arrest" as there is a response option for "no." The NPI team was able to link approximately half of the records missing data in the "subject arrested" field to the arrest data.

Finally, the APD's use of force data includes several injury-related fields with considerable missing data and logical inconsistencies. There are slight differences in the injury questions across Tier One, Two, and Three use of force summary reports. Each tier report includes two key questions: (1) whether the subject was injured (yes/no), and (2) the nature of the subject's injury (free text field). At the subject level, 75.6% of individuals who had force used against them (2,858 out of 3,783) do not have a valid entry for whether a subject was injured. Of those 2,858 individuals, 94% were also missing the nature of the subject injury.[24]

In other cases, the injury-related fields contradicted one another.[25] For example, for 111 individuals, the yes/no "subject injured" field indicated no injury, but an injury description was provided. In some instances, this may be due to information being provided in the injury description field that was related to a pre-existing injury rather than an injury associated with the current use of force.[26] Finally, data fields for officer and subject injury are not linked to the force type used. In cases where an officer used

---

[23] The response options for the alcohol and drug impairment data fields include: no, suspected, yes, and unknown. Despite the "unknown" response field, the percentage of missing data remains very high for these two variables.

[24] When there was an entry for nature of the subject injury, all but one said "unknown."

[25] Another example of inconsistency in injury-related data fields was found among the sampled cases. Within a single incident, the same injury nature was listed for multiple subjects "OC spray in the eye" but in the yes/no "subject injured" data field, yes was selected for some and no was selected for others.

[26] Specifically, the use of force reports include fields prompting the description of a subject's pre-existing injury and the nature of that injury. However, many reports documented pre-existing injuries within the "subject injury" description field reserved to describe injuries resulting from use of force. This created inconsistencies in the data reported.

a single type of force, reported injuries can logically be linked to that force type. However, in cases where more than one type of force was used, the data do not permit analyses related to the types of force that led to officer or subject injuries.

Summary

Despite multiple attempts to correct issues with the APD use of force data, the combination of these problems does not allow for valid and reliable analyses typically found in a use of force study, including:

1. Analyses of the type(s) of force used, including the effectiveness of each force action.
2. Analyses of the differences in use of force patterns across organizational units or geographic areas.
3. Analyses of types of force, injuries, or geographic patterns of use of force across racial/ethnic groups.
4. Analyses predicting the likelihood and severity of injuries to the officer or subject during use of force incidents.

Correcting the problems in APD's previously collected data is both time-intensive and cost-prohibitive. It would involve reading each report narrative to complete missing data or clarify contradictory data, if possible. Therefore, the NPI team proceeded with the best available information and limited analyses to those that could be conducted with confidence using these data. These analyses include: (1) time series analyses examining the trends in use of force counts, (2) the calculation of use of force disparity ratios by race/ethnicity based on benchmark comparisons, and (3) multivariate analyses predicting the likelihood of an arrest resulting in a use of force. The results of these analyses provide baseline measures against which the APD can compare future years of data.

STATISTICAL ANALYSES

To examine patterns and trends in APD enforcement data, four statistical analyses are conducted: (1) descriptive analyses, (2) time series analyses, (3) benchmark comparisons, and (4) multivariate analyses. These statistical techniques, their limitations, and the appropriate interpretation of their findings are described below.

Descriptive Analyses

To understand police enforcement actions, the first step is to describe the data available, and examine the patterns and trends of these data. Descriptive analysis is a fundamental component of data analysis that involves summarizing and presenting outcome count data. These analyses provide a clear and concise overview of key characteristics and patterns within a dataset, allowing analysts to gain insights into the data's central tendencies, variability, and distribution (Witte & Witte, 2015). Bivariate analyses or crosstabulations are a type of descriptive analysis examining the association between two variables (e.g., race and use of force). Descriptive analyses provide a critical basis for understanding basic patterns and distributions in the data and offer an initial assessment of the general trends and potential correlations between the predictor and outcome variables before primary analytical techniques are employed. Descriptive analyses are often limited in scope, cannot be used to explain or predict trends, and provide limited implications regarding findings. Thus, they are typically used as a precursor to more complex statistical techniques and illuminate appropriate methodological approaches (Witte & Witte, 2015).

Interrupted Time Series Analyses

It is important to consider how patterns and trends in police enforcement actions fluctuate over time. Interrupted time series analyses are considered one of the strongest quasi-experimental designs to determine whether the timing of a relevant intervention (e.g., police training or policy change) or a seminal event of interest (e.g., an arrest or use of force incident of public interest) corresponds with a significant shift in count outcomes, such as arrests or use of force counts (see Hudson et al., 2019). The key feature of interrupted time series analysis is the collection of data on the frequency of a specific outcome aggregated at regular time points before and after the intervention or event. It is considered best practice to aggregate the data into a monthly[27] time series format with a sufficiently long pre-intervention period (i.e., at least two years of monthly data) that allows researchers to determine whether there is a statistically significant change in the outcome immediately following the intervention, while also accounting for any pre-existing trends or patterns in the data. Time series analyses also require a sufficiently long post-period, which ranges from a minimum of seven to 12 months.[28]

Across the six-year period examined, numerous seminal events occurred that may have impacted – or disrupted – preexisting patterns in crime and police enforcement in Aurora. The NPI team identified ten such significant events to consider when analyzing trends in crime and APD enforcement activities over time, which are listed in **Table 2.7**

[27] Traditionally, monthly event counts are preferred over weekly event counts because the data are more stable and consistent across multiple years of observations.
[28] CrimeSolutions.gov is a warehouse for the National Institute of Justice's evidence-based strategies and programs, which experts review and score for their scientific merit. For these programs, any strategy that has a follow-up period of less than 7-months is gauged as a 'short term' program, while a one-year follow-up is required to be considered a long-term program. Consistent with this framework, we obtain 7 to 12 months post-period for time series assessments to be consistent with rigorous evaluations. See also Corsaro (2022).

below. These ten specific events of interest served as intervention points in the time series analyses presented in Section 3.

Table 2.7. Seminal Events and Dates Examined with Interrupted Time Series (ITS) Analyses

Seminal Event	Event Date
1. Death of Elijah McClain	August 2019
2. Colorado Executive Order declaring COVID-19 Disaster Emergency	April 2020[29]
3. Officer-involved death of George Floyd in Minneapolis	May 2020
4. Enactment of Colorado SB 20-217: Enhance Law Enforcement Integrity	July 2020
5. AG launches pattern or practice investigation	August 2020
6. Independent Review Panel report released	February 2021
7. Indictment of officers involved in McClain death	September 2021
8. City enters into Consent Decree	November 2021
9. APD Chief Vanessa Wilson terminated	April 2022
10. Interim APD Chief Dan Oates hired	June 2022

Benchmark Comparisons

Benchmarking analyses are often used to examine racial disparities in policing outcomes by comparing data against established "benchmarks" to assess fairness and equity in law enforcement practices. Benchmarking provides a standardized basis for evaluating disparities by comparing outcomes across racial or ethnic groups using an external data source to represent the "expected" population for that outcome. For example, to determine racial disparities in arrests, the percentage of the arrestee population who are Black is compared to the percentage of the benchmark population who are Black. The estimated "at risk" benchmark population that is selected drives the results. Studies have consistently demonstrated that the use of different benchmark populations can result in dramatically different findings. Therefore, it is critical to know and understand the strengths and limitations of the benchmark population being used. All benchmarks have limitations and vary in the extent to which they accurately estimate the population of similarly situated individuals "at risk" of police enforcement actions, assuming no bias exists (Engel & Calnon, 2004; PERF, 2021; Tillyer et al., 2010).

For benchmark analysis, the groups are compared in the frequency with which they experience a particular outcome (usually calculated as a rate), using some scaling factor

[29] The Colorado Governor's Executive Order occurred on March 10, 2020. Because the interrupted time series analysis requires monthly data, April 1, 2020 is used in all models to distinguish the pre- and post-COVID onset as the first full month with Executive Order in effect. This may slightly underestimate its effect.

(such as the underlying population). If certain groups are found to experience significantly higher rates than expected based on their underlying risk set, this is typically interpreted as evidence of disparity. Conducting benchmark analyses promotes transparency and accountability and has been applied across various outcomes in criminal justice to highlight areas where disparities are more pronounced. It can be used by policymakers and law enforcement executives to provide context, help guide reform efforts, and monitor the impacts of reform-related changes over time. For this report, benchmarking analysis is employed to examine racial/ethnic disparities in arrests and use of force. The general description of the procedure used by the NPI team below applies to both arrests and use of force analyses.

The most widely used external benchmark is the residential population, which compares the frequency of an outcome (e.g., arrest) by racial group to their representation in the residential population. Although intuitive, this methodology has been routinely demonstrated as flawed in its ability to identify and quantify racial disparities in law enforcement outcomes (Alpert et al., 2004; Fridell, 2004; Smith et al., 2019). This is because not all people who reside in a city or neighborhood have the same "risk" of police enforcement activity. For example, the risk of being arrested is influenced by many factors – including involvement in criminal activity – which may not be evenly distributed across the residential population. Census data do not measure the types of characteristics shown by research to put individuals at risk of experiencing force, including several legally relevant behaviors including subjects' resistance, presence of a weapon, and criminal behavior (Engel et al., 2000; Garner et al., 2002; Morgan et al., 2020). Using the residential population as a comparison benchmark does not include any accounting of the likelihood or risk of police enforcement activity, and, therefore, is one of the weakest benchmark comparisons. Also note that benchmark analyses (regardless of the comparison data source) lack the depth to explain the reasons behind any reported disparities by failing to consider the complex factors potentially contributing to differential outcomes across racial and ethnic groups.

The NPI team compares non-census-derived benchmarks that better approximate the risk of contact with police that could result in enforcement action to the percentage of racial/ethnic groups that receive police enforcement actions. These include the percentage of racial/ethnic groups among the following comparison data sources: (1) individuals issued criminal summonses, (2) arrested individuals (all offenses, Part I only, and Part I violent only), and (3) crime suspects as reported to the police (all suspects, Part I suspects only, and Part I violent suspects only).

Most individuals who experience use of force are arrested (Davis et al., 2018; Garner et al., 2018; Hickman et al., 2008), making arrest data a viable proxy measure for assessing risk of use of force. However, if there is police bias in who is arrested, then using arrest data to approximate the expected racial/ethnic percentages of those who experience force violates the assumption that no bias exists and may underestimate disparity (Cesario et al., 2019; Geller et al., 2021; Knox et al., 2020a, 2020b; Knox & Mummolo, 2020). Furthermore, not all use of force situations result in arrests. This is another limitation of using arrest data as a benchmark for measuring racial/ethnic disparities in use of force.

Criminal suspect data is another benchmark used to approximate risk of police enforcement contacts. This information is collected by the police through crime reports. Here the information is based on community members' experiences and descriptions (Ridgeway & MacDonald, 2010; Smith et al., 2022). While this addresses one of the limitations of arrest benchmarks (potential officer bias in arrests), the criminal suspect-based benchmarks may reflect the likelihood of community members reporting certain types of crimes more than others (e.g., violent crimes more so than property crimes) (Klinger & Bridges, 1997), which may or may not be related to the likelihood of use of force. Likewise, reported crimes may themselves be biased against offenders of certain racial/ethnic groups based on the willingness of community members to report victimization.

The research available regarding the validity of different benchmarks and the factors that influence police behavior suggests that criminal suspect-based benchmarks are stronger approximations of the population "at risk" of being arrested or having force used compared to other benchmarks, while residential census data is widely considered an unreliable and invalid comparison measure (Alpert et al., 2004; Fridell, 2004; Geller et al., 2021; Smith et al., 2019; Smith et al., 2022).

To examine racial disparities in arrest and use of force, the NPI team calculates *disparity ratios*, a useful and easily interpretable technique for comparing groups who experienced force (or arrest) to those groups at risk for force relative to the non-Hispanic, White population (Smith et al., 2019). The calculation of the disparity ratio is a two-step process. First, the disproportionality index (DI) is calculated by dividing a racial group's representation in use of force incidents (or arrests) by the same group's representation in the comparison benchmark (e.g., suspect population). The result of this calculation measures within-group differences. Values greater than one indicate that the group experienced police enforcement actions more often than would be expected based on their representation in the benchmark. In contrast, a value of less than one indicates they experienced enforcement actions less often than expected based on the same benchmark. Second, the disparity ratio can be calculated to measure between-group differences by dividing the DI of the minority group by the DI of the majority group. A disparity ratio greater than one suggests that Black or Hispanic individuals were more likely than their White counterparts to experience police enforcement actions based on the benchmark used, whereas a disparity ratio less than one indicates the opposite. While disparity ratios are a useful method of estimating the size of disparities, there is no threshold value at which disparity can be attributed to racial *bias* (Fridell, 2004; Geller et al., 2021). For example, disparity ratios greater than one do not imply the existence of police bias; likewise, disparity ratios equal to one do not imply the absence of bias. Only the presence of disparities can be calculated with benchmark analyses, not the presence of bias.

Previous research shows that benchmark comparisons based on population statistics nearly always show racial/ethnic disparities in use of force, while benchmarks based on arrests or reported crime suspects show reduced or no racial/ethnic disparities (Brown et al., 2022; Cesario et al., 2019; Fryer, 2019; Geller et al., 2021; Ross et al., 2020, Tregle et al., 2019). Despite its limitations, the NPI team includes benchmark comparisons based on the 2020 US Census, along with non-census benchmarking, for two narrow purposes. First, these analyses provide a baseline of how different racial/ethnic groups experience enforcement actions. Second, the comparison of disparity ratios across a variety of benchmarks helps to determine the validity of the analytical technique for representing the population at risk police enforcement actions.

Multivariate Logistic Regression Analyses

Multivariate regression modeling is a statistical technique that creates a mathematical equation that considers the influence of multiple variables on an outcome. For example, to understand the impact of subjects' race on the likelihood of having forced used, a multivariate regression model can estimate the impact of various factors (other than race) on the likelihood of use of force of persons who are arrested. Here, the population (arrestees) is known (through arrest reports); likewise, whether force is used during the arrest encounter is also known (through use of force reports). The mathematical equation generated for regression modeling helps to predict or understand how changes across multiple factors might affect the likelihood of police use of force.

While they are different analytical techniques, benchmark analyses are complemented by multivariate regression analysis. Unlike benchmark analyses, there is no need to make comparisons to an estimated benchmark population because both populations (arrestees and those who had force used against them) are known and used in regression analyses. While benchmarking may help identify disparities, multivariate regression modeling helps uncover those complex underlying factors contributing to different outcomes. Multivariate regression provides a more nuanced understanding by considering multiple variables simultaneously, offering insights into the interplay of factors contributing to racial disparities. For example, to know if Black subjects are more likely than White subjects to have force used against them during arrest situations, it is important to simultaneously consider other factors (e.g., other characteristics of the person, situation, and neighborhood) that may also impact if force is used. Instead of focusing on just one component that may affect officer decision-making, (e.g., subject's race), multivariate regression quantifies the impact of multiple factors simultaneously and estimates how confident we can be that these results are not due to random chance.

Multivariate logistic regressions are typically employed to investigate complex relationships between multiple variables and assess their collective influence on binary outcomes, such as the decision to use force (Long, 1997; Witte & Witte, 2015).[30] Multivariate logistic regression techniques quantify the strength and direction of associations between various factors and the likelihood of use of force, while controlling for potential confounding influences (Hanushek & Jackson, 1977; Meyers et al., 2016). The key factors (i.e., independent variables) typically included in analyses to predict use of force within arrests include: (1) legal characteristics (e.g., outstanding warrants, type of criminal charges, presence of weapon, suspects resistance, etc.), (2) incident or situational characteristics (e.g., incident location, day, time, presence of bystanders, etc.), and (3) subject's demographic characteristics (e.g., age, race/ethnicity, gender).

Within logistic regression models, the estimated effects of the different variables are typically expressed as *odds ratios*, which indicate how strongly those factors are related to the outcome using a standardized scale.[31] An odds ratio greater than one indicates the variable is associated with higher odds of the event occurring, while an odds ratio less than one suggests association with lower odds of that occurrence. The standard guidance regarding the size of odds ratios suggests that odds ratios less than 1.5 are substantively small, 1.5 to 2.5 are medium, and 2.6 or greater are substantively large (Chen & Chen, 2010). The reported regression results also include point estimates (measuring the average change in the outcome when a factor changes) and significance values indicating our confidence in the results for the regression models.[32]

Predicted probability analyses precisely estimate how independent variables in the regression models impact a specific outcome. The predicted probability indicates the likelihood of an event (e.g., the chance of force during an arrest) while controlling for the rest of the factors in the model. These estimation methods reveal what factors are statistically associated with the outcome (e.g., use of force) and after considering everything else included the model, the exact chances of that event occurring.

[30] Multilevel modeling is appropriate for data collected across different units of aggregation and produces unbiased estimates at each of the analysis levels (Raudenbush & Bryk, 2002). Importantly, the arrest data include variables that cross units of analysis (i.e., nested data). Arrest incidents are nested within officers, which are nested within geographic units. The NPI team attempted multi-level modeling, however the lack of statistical power and reliability concerns with geographic mapping coordinates (derived from various sources), as well as the proportion of use of force cases that occurred outside of Aurora's boundaries, limited the capacity to conduct multilevel modeling.

[31] The odds ratio is the exponentiated coefficient given the logarithmic distribution used in logistic regression models.

[32] Statistical significance is expressed as a p-value of 95% confidence intervals, which are the standard of scientific rigor required in most social sciences (Betensky, 2019)

The major limitation in multivariate regression models is that the results only measure variables included in the analysis. Unmeasured or unincluded variables can potentially bias estimates and results. This is referred to as *model misspecification* or *omitted variable bias* (Hanushek & Jackson, 1977; Jung et al., 2018; Marvell & Moody, 1996). This is an important limitation because no single data form or report can reliably quantify all relevant information regarding officer decision-making. When interpreting the multivariate regression results, the NPI team takes care to note what the models mean and what they do not mean (based on omitted variables, where they exist). As noted previously, the APD arrest data do not include several potential explanatory factors of use of force, including measures of resistance, impairment, and weapon presence.[33] The exclusion of these factors from the statistical models severely limits our confidence in the validity of the findings.

Summary

Each type of statistical analysis has strengths and limitations that should be considered when interpreting the findings. Combining statistical approaches allows for more comprehensive policy recommendations, by understanding partners and trends over time (descriptive analyses, interrupted time series analyses), addressing observed disparities (benchmark analyses) and identifying possible contextual factors that contribute to police enforcement actions (multivariate regression analysis). Although benchmarking is valuable for identifying and quantifying racial disparities, multivariate regression analyses supports the examination of the complex interplay of contributing factors. A holistic approach incorporating all statistical methods can offer a more comprehensive understanding of racial disparities in policing outcomes and inform effective policy interventions.

SECTION SUMMARY

The NPI team analyzed several official APD data sources using multiple statistical techniques to understand patterns and trends in APD enforcement for the period of January 1, 2017 – December 31, 2022. The primary data sources used were (1) criminal offenses (incidents and suspects), (2) criminal summonses, (3) arrests, and (4) use of force.

---

[33] It is important to note that arrest data include charges, and criminal suspects may be charged with public intoxication, operating a motor vehicle under the influence, disorderly conduct, and resisting arrest (among others). However, these are not systematically available in all arrest reports, but rather would only represent when an officer charges the individual with an offense within these various categories. Since these situational characteristics are only collected when arrest charges occur, they are not included in any systematic analysis.

The NPI team used the criminal offense data to (1) examine trends in criminal *incidents* over time as context for the analyses of trends in APD enforcement activities (n=33,495 incidents) and (2) facilitate benchmark comparisons between racial/ethnic percentages of those who were arrested or experienced force and racial/ethnic percentages of known criminal *suspects* (n=35,889). The NPI team used the data on 20,922 individuals issued criminal summonses to examine trends over time and to facilitate benchmark comparisons for those who experienced force. The NPI team used data for 44,954 arrested individuals for three purposes: (1) to examine arrest trends over time and across race/ethnicity, (2) to facilitate benchmark comparisons for those who experienced force, and (3) to understand the factors that predict whether arrests result in force. The NPI team analyzed data for 3,783 individuals who had force used against them during 3,518 use of force incidents to examine trends over time and across race/ethnicity and to facilitate benchmark comparisons using force data as the numerator. Unfortunately, data limitations restricted the team's ability to complete an in-depth analysis of APD's use of force data.

The statistical analyses conducted with these data include basic descriptive analyses, time series analyses, benchmark analyses, and multivariate statistical modeling. *Descriptive analyses* provide researchers with a foundation for further data analysis, hypothesis testing, and decision-making by offering insights into the data's central tendencies, variability, and distribution. *Time series analyses* test whether the timing of a relevant intervention or seminal event corresponds with a significant shift in counts of outcomes of interest (e.g., crime, arrests, use of force), controlling for time-varying factors. The NPI team examined the impact of ten seminal events that occurred nationally and locally during the study period and their impact on monthly counts of different APD outcomes.

Benchmark analyses examine patterns of racial disparity by comparing the percentage of racial/ethnic groups arrested or experiencing force with the percentage of racial/ethnic groups' representation in comparison data sources (i.e., the benchmarks) that attempt to approximate the risk of use of force or arrest. *Multivariate analyses* simultaneously consider different factors and estimate significant predictors of the likelihood of force being used during arrests. This allows for estimating the individual impact of race/ethnicity while accounting for other key factors that may impact whether officers use force.

In summary, the NPI team conducted a series of statistical analyses to better understand APD's enforcement activities from 2017 to 2022. It is important to consider the results collectively, while considering the strengths and weaknesses of the data sources and statistical techniques used. Further, the findings should be interpreted through an understanding of the context in which enforcement decisions are made by officers. The findings that emerge using multiple approaches can then be used as baseline measures to examine the impact of police reforms implemented by the APD over time.

SECTION 3: CRIME, CRIMINAL SUMMONSES, ARRESTS, AND USES OF FORCE – TRENDS OVER TIME

In this section, the NPI team examines patterns and trends reported by the APD of the following: (1) criminal incidents, (2) criminal summons, (3) arrests, and (4) uses of force. These data points and their sources are described in Section 2. The NPI team examines these data sources for a six-year period, from Jan 1, 2017–Dec 31, 2022.

One of the analytical techniques used to explore the patterns and trends of these data is interrupted time series analysis, specifically to determine the impact of seminal events that occurred during the six-year period (list of seminal events provided in Section 2). The purpose is to determine if these data should be examined as a continuous, uninterrupted data source or if a particular event or series of events changed the trajectory or pattern of criminal or police activity. Note that many of the seminal events of interest were relatively close to one another in time, making interpreting the findings challenging. Nevertheless, the time series analyses exploring crime and police activity in Aurora – including individual examinations by racial/ethnic groups – demonstrate that one event, in particular, had an abrupt and disruptive influence on otherwise pre-existing and stable patterns of activity. Specifically, the onset of the COVID-19 pandemic produced an unmistakable, immediate impact on the counts of certain types of police activity. The results from these analyses are presented in detail below. **Our overall takeaway is that the APD enforcement data from the six-year period of interest should be analyzed separately as two distinct and comparative periods: (1) Pre-COVID (Jan 2017–Mar 2020), and (2) Post-COVID (Apr 2020–Dec 2022).**

APD REPORTED CRIMINAL INCIDENTS, JAN 2017–DEC 2022

From January 1, 2017, to December 31, 2022, the NPI team received data for 49,173 criminal offenses resulting from 33,495 incidents involving 35,889 individuals. This section examines trends in the 33,495 criminal incidents over time to understand patterns in APD enforcement activities. **Figure 3.1** displays the annual counts of overall criminal incidents, which includes any incident with at least one criminal offense (felony or misdemeanor) for the six-year study period (2017–2022). Likewise, **Figure 3.2** displays the annual counts of all reported criminal incidents with at least one Part I offense (burglary, larceny, motor vehicle theft, and Part I violent offenses), along with the subset of reported criminal incidents with at least one Part I violent offense (aggravated assault, robbery, rape, and murder).

Figures 3.1 and 3.2 below show that criminal incidents in the City of Aurora have steadily increased since 2017. Overall, there was a 20% increase in total criminal incidents from 2017 to 2022, including a 44% increase in total Part I offenses and an 82% increase in Part I violent offenses across this six-year period.

Figure 3.1. Annual Counts of Criminal Incidents Reported to APD, 2017–2022 (n = 33,495)

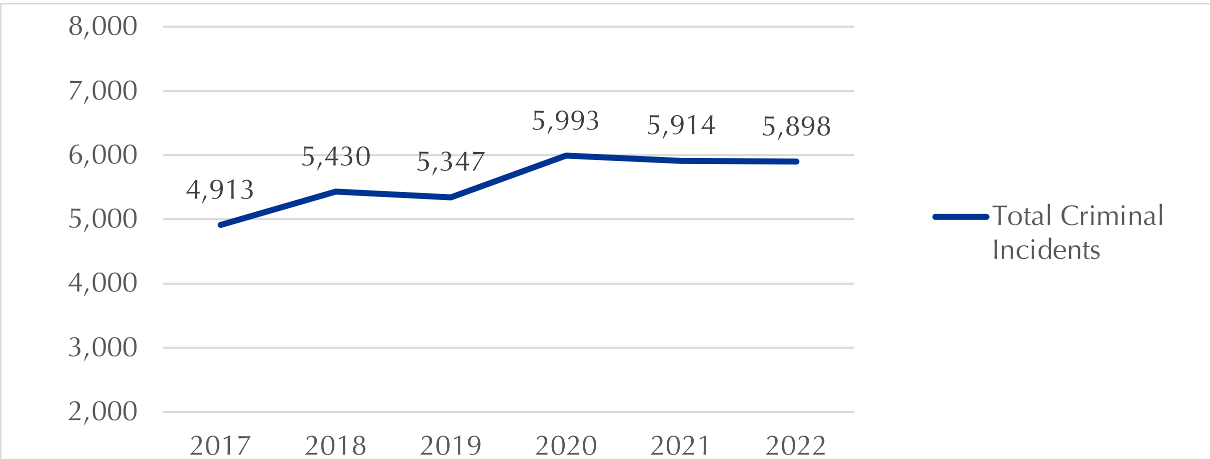
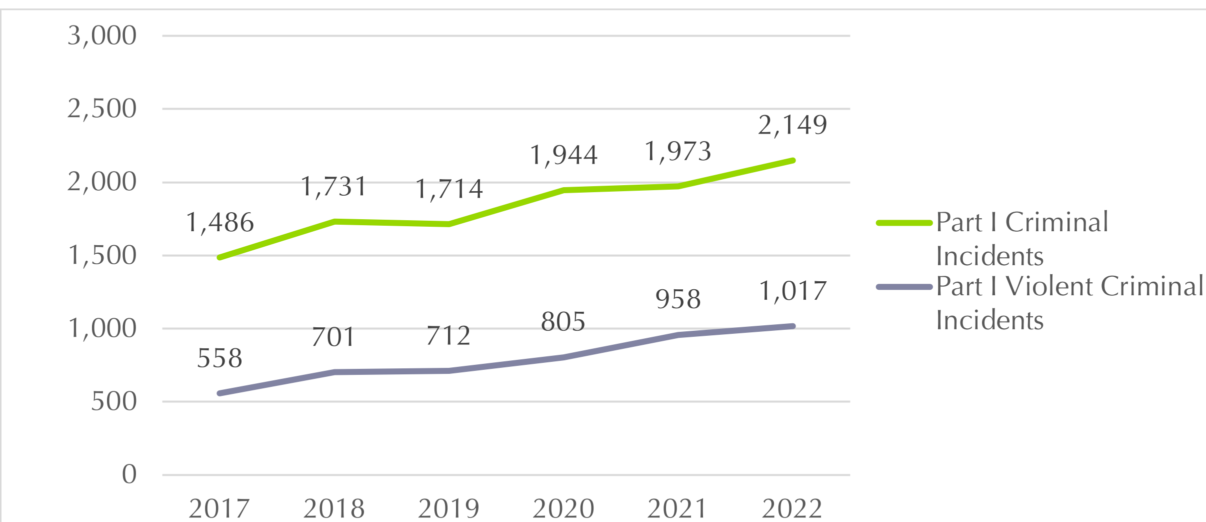


Figure 3.2. Annual Counts of Part I and Part I Violent Criminal Incidents Reported to APD, 2017–2022



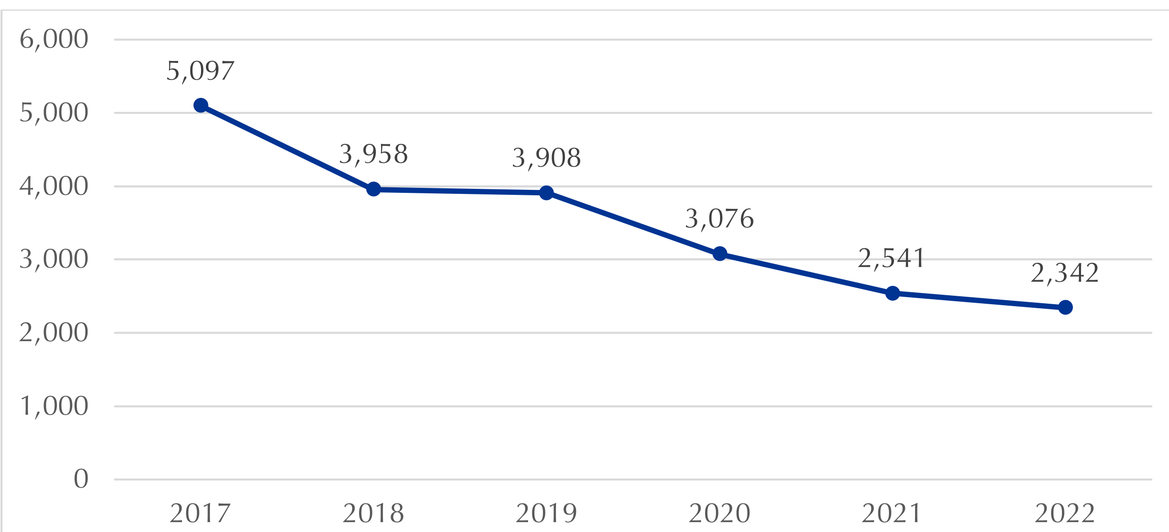
To provide additional context, the NPI team conducted supplemental time series analyses using ten different intervention dates of interest to examine trends in reported crime after seminal events. None of the time series analyses demonstrated a statistically significant shift for reported criminal incidents in Aurora (results available upon request). Instead, criminal incidents appeared to follow a consistent upward trend that was not significantly altered (reduced or accelerated) by the seminal events examined, including the onset of the COVID-19 pandemic post-March 2020.

APD CRIMINAL SUMMONSES, JAN 2017–DEC 2022

While arrests are the primary source of contact where APD charges individuals with criminal offenses, they are not the only type of enforcement contact between APD officers and members of the public. APD officers also issue criminal summonses (i.e., a summons to appear in court where the person is not taken into custody). In total, APD officers issued 20,922 criminal summonses from 2017 to 2022, which are the focus of analyses in this section.

Figure 3.3 shows the trends in criminal summonses from Jan 1, 2017–Dec 31, 2022. Analyses of criminal summonses highlight a linear decline across the entire six-year period and a specific post-COVID decline. Criminal summonses peaked with over 5,000 issued at the beginning of the period examined (2017) but declined significantly to 3,958 and 3,908 in 2018 and 2019, respectively. This decline continued to just over 3,000 summonses in 2020 and lowered further to just over 2,500 and 2,300 in 2021 and 2022, respectively. Over the six-year period, the number of criminal summonses issued by the APD declined by 54.1%.

Figure 3.3. Annual Counts of APD Criminal Summonses, 2017–2022 (n = 20,922)



As shown in **Table 3.1**, the racial and ethnic distribution of individuals who received criminal summonses was consistent from 2017 to 2022, with a relatively uniform distribution across the entire period. The reduction in criminal summonses did not impact the distribution across racial and ethnic groups. Criminal summonses declined by greater than 50% from 2017 to 2022 for White (-54.4%), Black (-50.3%), and Hispanic (-50.7%) individuals.

Table 3.1. Annual Counts of APD Criminal Summonses by Race/Ethnicity, 2017–2022

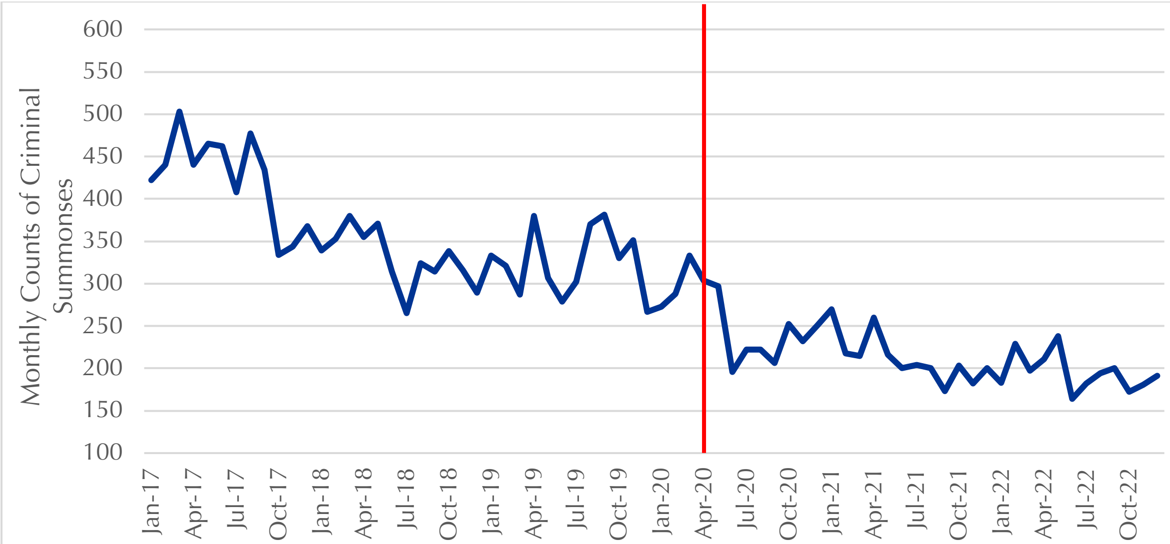
	Total 2017-2022	2017	2018	2019	2020	2021	2022	% Change 2017-2022
White	6,990	1,667	1,329	1,227	1,089	918	760	-54.4%
Black	7,590	1,758	1,393	1,466	1,181	919	873	-50.3%
Hispanic	5,342	1,288	998	1,060	723	638	635	-50.7%

NOTE: This table excludes 1,000 criminal summonses issued to individuals of “other” (n=597) or “unknown” (n=403) races across the six-year period.

As with our examination of reported crime, the NPI team considered whether the trends in issuing criminal summonses were altered by any seminal events using interrupted time series analyses. The monthly trends are graphically displayed in **Figure 3.4**, demonstrating the downward linear trend in criminal summonses. Again, ten dates of seminal events that could potentially impact police enforcement activities (see Section 2) were examined for changes in the monthly counts of total criminal summonses and by racial/ethnic group.[34]

[34] The criminal summonses time series also required the inclusion of a linear trend control variable to account for the constant linear decline/shift in the monthly count of criminal summonses for the entire time series period.

Figure 3.4. Monthly Counts of APD Criminal Summonses, 2017–2022 (n = 20,922)



The interrupted time series analyses demonstrate that the sole seminal event that significantly impacted the issuing of criminal summonses was the onset of the COVID-19 pandemic (post-March 2020). While the linear trend of criminal summonses counts by month had been steadily decreasing, the sudden impact related to COVID-19 accelerated this decline, over and above what would be expected from the pre-existing trends and seasonal variations. As reported in **Table 3.2**, post-March 2020, there was a statistically significant reduction in total criminal summonses of 11.2%, which can be attributed to the shift in summonses from April 2020 through December 2022.

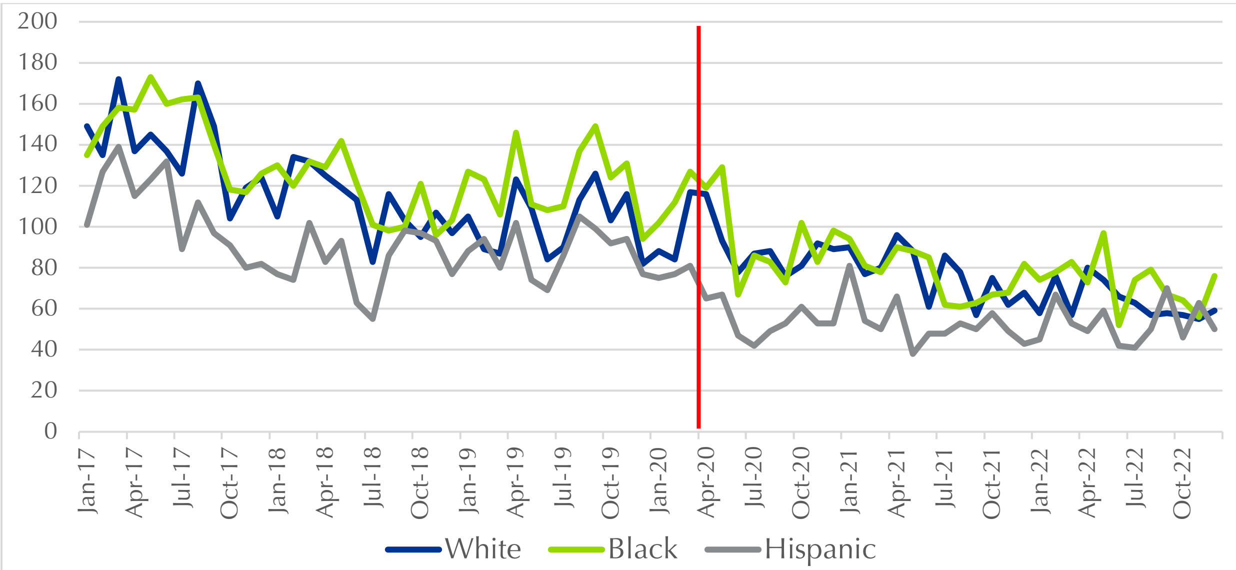
Table 3.2. Interrupted Time Series Analyses for APD Criminal Summonses, 2017-2022

	Total Summonses	Standard Error	Exp(B)-1
	B		
Intercept	6.05*	0.038	--
Post-COVID	-0.119*	0.052	-0.112

NOTE: All regression models include February – December monthly dummy variables (included in models but excluded from tables for parsimony) and a linear-trend variable. *p < 0.05

Figure 3.5 shows that the number of Black, White, and Hispanic individuals who received criminal summonses declined linearly from 2017 to 2022. The trend analyses indicate a reduction in criminal summonses that transcended the race/ethnicity of suspects. However, time series analyses modeling the impact of the COVID-19 pandemic on changes in criminal summonses across race/ethnicity indicate that there was not the same abrupt, permanent shift post-March 2020 period for White suspects relative to Black and Hispanic suspects. While the reductions over time were similar across all three racial and ethnic groups, the COVID-19 shift only impacted reductions in criminal summonses issued to Black and Hispanic individuals. Unfortunately, the character and quality of the criminal summons data did not allow the NPI team a reasonable method for examining the frequency and impact of different types of summonses.[35]

Figure 3.5. Monthly Counts of APD Criminal Summonses by Race and Ethnicity, 2017–2022 (n = 20,922)



APD ARRESTS, JAN 2017–DEC 2022

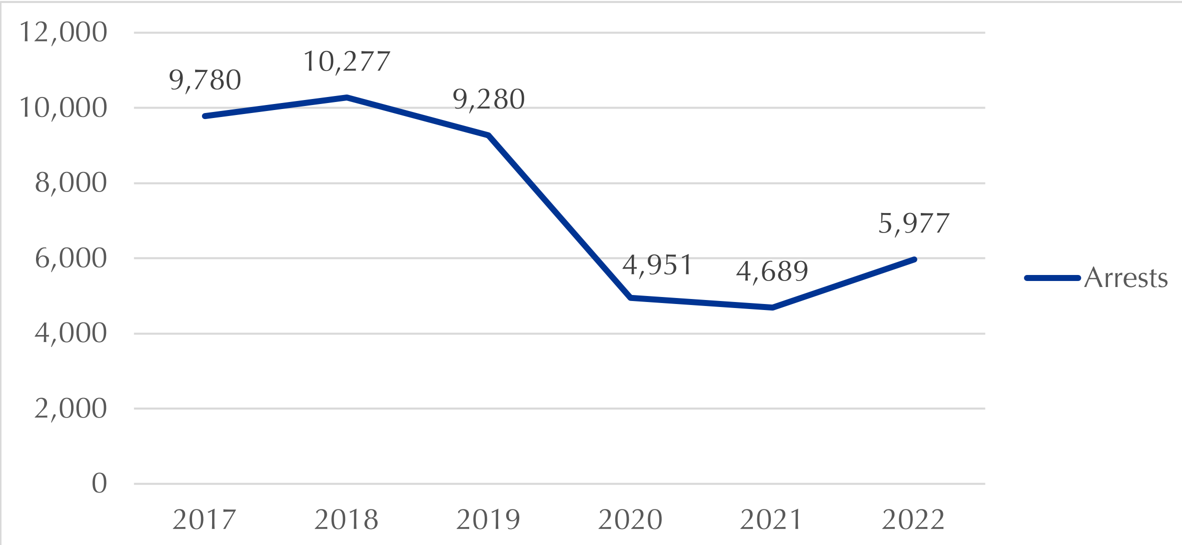
As noted in Section 2, individual arrests are counted at the incident-person level. For example, a single person could be arrested multiple times over the six-year period. Each arrest incident involving the same person is counted as an independent arrest in these situations. In addition, multiple individuals could be arrested in a single incident (i.e., on the same date/time at the same location, involving the same circumstances). In these situations, each individual is counted as an independent arrest. Using this definition, across the APD, officers arrested 44,954 individuals during the six-year study period.

Figure 3.6 graphically displays the distribution of arrests by year. As shown, the number of arrested individuals was roughly stable from 2017–2019 (avg = 9,779 per year) before a steep decline from 2020–2022 (avg = 5,200 per year). When considered as distinct three-year periods (2017–2019, compared to 2020–2022), there is a nearly 47% decline

[35] The criminal summons data included 1,039 different manual entries or "string" text fields; examination of these data fields (which would have required hand coding) was beyond the scope of NPI's engagement.

in overall arrests in the post-2020 period relative to the pre-2020 period.[36] This arrest decline occurred simultaneously with a significant increase in reported crime in the City of Aurora across the full six-year period, including an 82% increase in violent incidents.

Figure 3.6. Annual Counts of APD Arrests, 2017–2022 (n = 44,954)



Further examination of APD arrest trends shows that the decline in arrests beginning in 2020 occurred across all racial/ethnic groups (see **Table 3.3**). Specifically, average yearly arrests of White individuals declined by 49.8% comparing the periods pre- and post-2020, followed closely by a 47.3% decline in arrests of Black individuals and a 41.2% decline in arrests of Hispanic individuals.

Table 3.3. Annual Counts of APD Arrests by Race/Ethnicity, 2017–2022

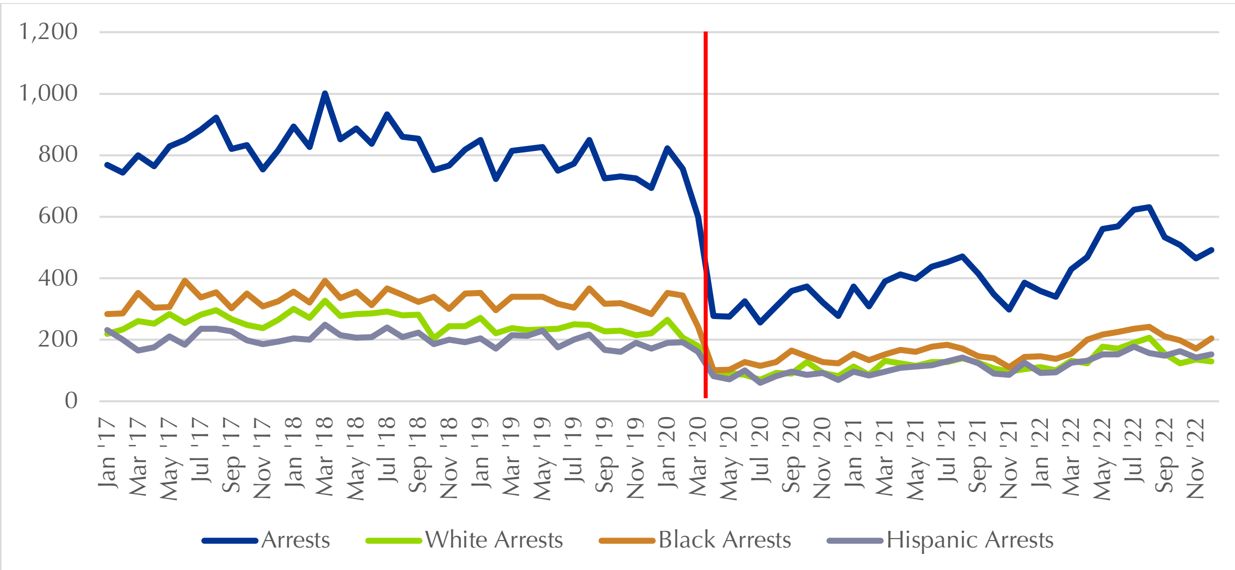
	Overall 2017-2022	2017	2018	2019	Avg 2017-2019	2020	2021	2022	Avg 2020-2022	% Change Between Avgs
White	13,838	3,098	3,289	2,826	3,071	1,470	1,398	1,757	1,542	-49.8%
Black	18,137	3,906	4,098	3,874	3,959	2,074	1,842	2,343	2,086	-47.3%
Hispanic	11,579	2,445	2,534	2,313	2,431	1,282	1,313	1,692	1,429	-41.2%

NOTE: 1,400 arrestees of "other" (n=1,069) or "unknown" (n=331) races are excluded from this table.

[36] Percentage change is a bivariate change only, meaning it examines the distinct shift in arrest counts between the two time periods but does not account for any seasonal fluctuations or trends in the data.

To better understand APD arrest patterns, the NPI team conducted interrupted time series analyses[37] on monthly arrest data. **Figure 3.7** below disaggregates the monthly arrest counts by suspects' race/ethnicity. As shown, the monthly bivariate trend change indicates that all arrests – across racial/ethnic groups – declined considerably in April 2020.

Figure 3.7. Monthly Counts of All APD Arrests, 2017–2022 (n = 44,954)



The interrupted time series analyses results in **Table 3.4** demonstrate that the total number of APD arrests was significantly influenced (i.e., abruptly reduced) by the onset and response to the COVID-19 pandemic in mid-March, resulting in a 49.8% reduction in arrests for the post-March 2020 period examined.

Table 3.4: Interrupted Time Series Analyses for APD Arrests, 2017–2022

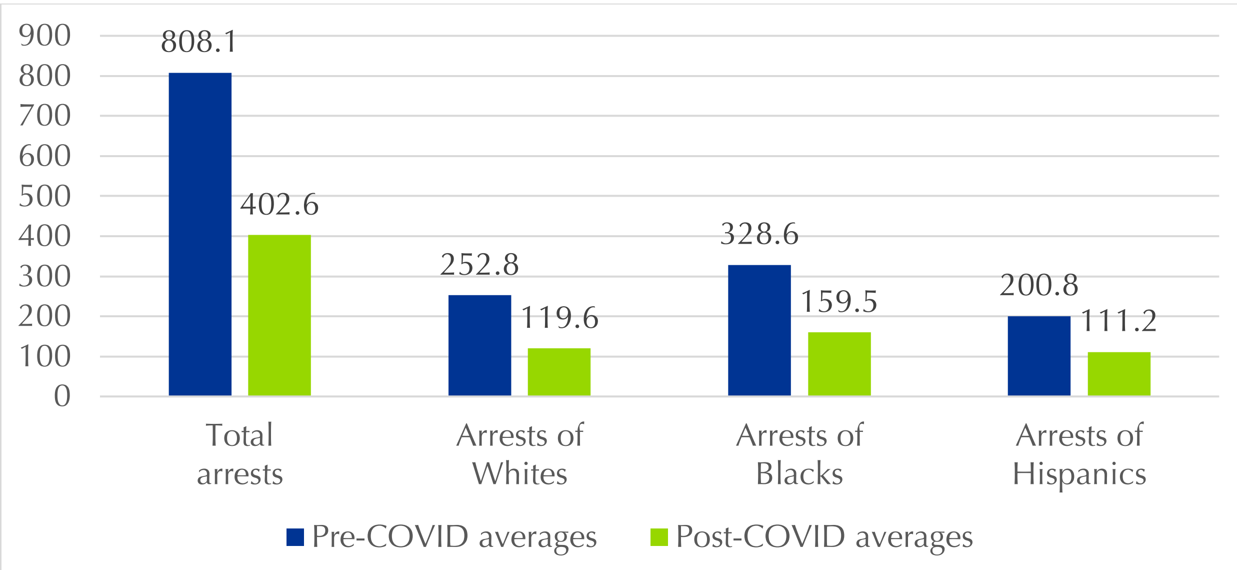
	Total Arrests	Standard Error	ExpB-1
	B		
Intercept	6.70*	0.030	--
Post-COVID	-0.690*	0.041	-0.498

+All regression models include February – December monthly dummy variables (included in models but excluded from tables for parsimony) *p <0.05

[37] The details of interrupted time series analyses are described in Section 2. Each statistical model included monthly dichotomous variables to account for seasonality, and robust standard errors to mute a degree of the presence of serial autocorrelation between residuals (to address that the time periods immediately preceding and following a specific period are correlated over time).

Additional time series analyses (available upon request) demonstrate that the post-March 2020 period reduction was significant across racial/ethnic groups. **Figure 3.8** displays these declines from the pre-COVID (Jan 2017–Mar 2020) to post-COVID (Apr 2020–Dec 2022) periods, where the pre-COVID average total arrests per month (808.1) was reduced by 50.2% to 402.6 arrests per month post-COVID. These analyses account for seasonal fluctuations in the data and are more accurate and precise estimates of change relative to the earlier bivariate percentage change in arrests. As shown, arrests of White individuals declined by 52.7%, arrests of Blacks individuals declined by 51.5%, and arrests of Hispanic individuals declined by 44.6% in the post-COVID period.[38]

Figure 3.8. Pre & Post-COVID Comparison of Average APD Arrests, Overall and by Race, 2017–2022 (n = 44,954)



A second statistically significant change in arrest patterns was observed when examining the seminal event of the murder of George Floyd in Minneapolis in late May 2020. Although APD arrests declined immediately following the onset of the pandemic (post-March 2020), as shown in **Table 3.5**, the independent variable that captures the post-May 2020 period (i.e., post-George Floyd period) saw a *rebound* and significant *increase* in total arrests by roughly 50% (i.e., an increase of half of the 50% decline in total arrests that occurred post-COVID). This statistically significant increase in arrests beginning in June 2020 was observed across all racial and ethnic groups. However, the increase for Black arrestees was slightly higher than that of other racial/ethnic groups (results available upon request).

[38] These percentages are calculated as follows: (value 2 - value 1)/value 1 x100. For example, for arrests of Blacks (159.5 – 328.6)/328.6 = -.5146 x100 = -51.46% change or 51.46% decrease.

Table 3.5. Interrupted Time Series Analyses for APD Arrests, 2017–2022

	Total Arrests	Standard Error	ExpB-1
	B		
Intercept	-1.08*	0.027	--
Post-Floyd	0.409*	0.040	[0.505]

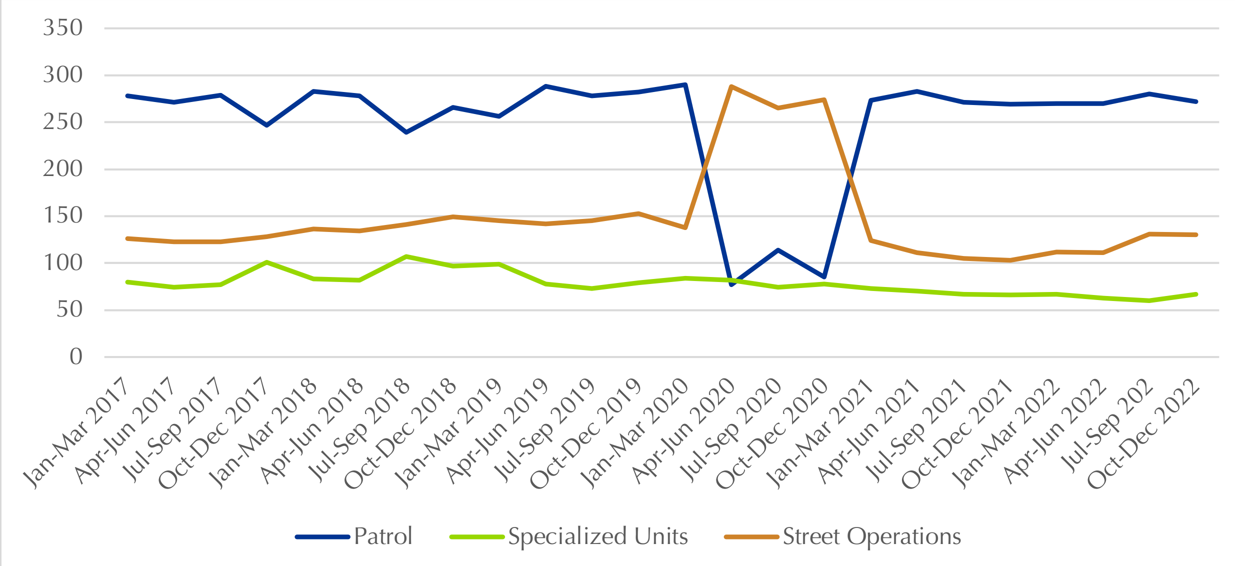
+All regression models include Feb – Dec monthly dichotomous variables (excluded from tables for parsimony) *p < 0.05

Examining Reductions in Arrests

Again, it is important to note that the observed significant reduction in APD arrests occurred during increased reported crime and violence. This suggests the change in arrest patterns observed was potentially a product of changes in APD's approaches to various operational issues (potentially associated with the COVID-19 pandemic or post-Floyd response) rather than a change in crime.

To examine these possibilities, the NPI team conducted simple trend comparisons of quarterly counts of personnel assigned to APD patrol operations during this period. As shown in **Figure 3.9**, outside of a brief period in 2020 where patrol personnel declined and street operations personnel increased, the quarterly counts of APD personnel demonstrate a relatively stable patrol operational force during this study period. The proportion of APD personnel assigned directly to patrol showed a similar pattern, with patrol comprising an average of 38% of the total personnel (outside of the divergence in 2020). Thus, the overall change in arrests was not associated with a significant shift in the number of sworn APD personnel. While there was a reduction in the number of personnel assigned to patrol in the 2020 COVID year, the number of overall sworn personnel was stable. Furthermore, the shifting of personnel assignments returned to a similar pre-COVID level in 2021 while the sustained decline in arrests continued. Combined, these patterns suggest the abrupt and sustained decline in arrests is not a product of APD personnel changes.

Figure 3.9. Quarterly Counts of APD Employees by Unit, 2017–2022



As a final step in understanding the decline in arrests, the NPI team examined arrest trends involving different offense charges, focusing on which charges declined the most in the post-March 2020 period. For any arrested individual, there may be one or more charges issued. **Figure 3.10** below shows the pre- and post-March 2020 (COVID-19 pandemic) percentage changes for arrest *charges* of interest. In summary, the bivariate descriptive analyses indicate the following patterns:

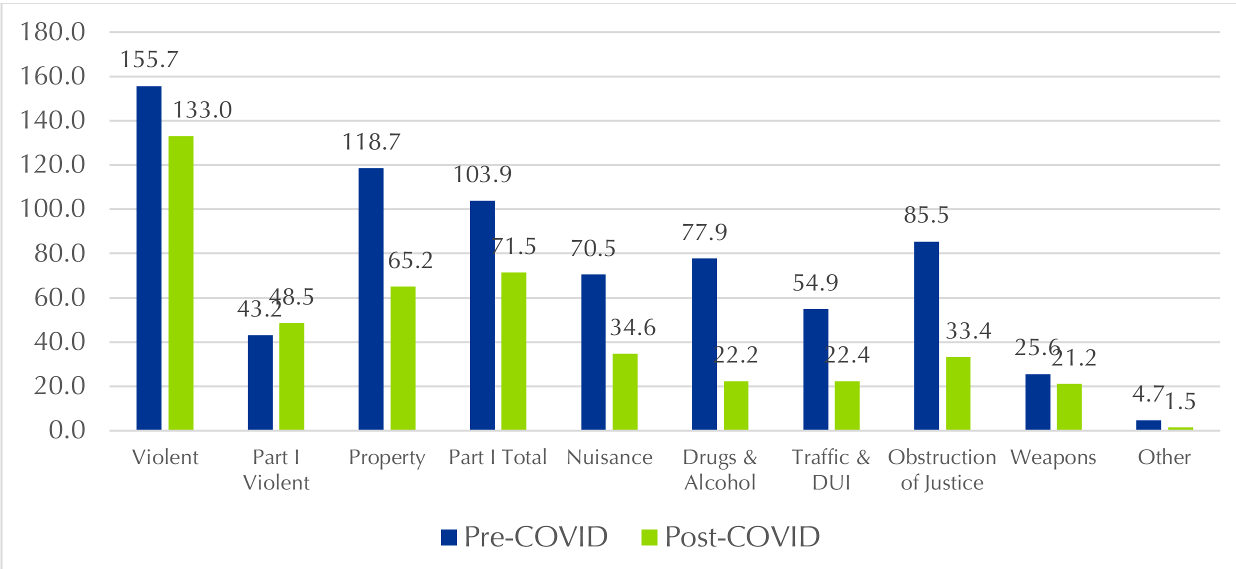
- A sizable and long-term post-March 2020 decline occurred in arrest charges involving drug and alcohol charges, traffic charges, obstruction of justice charges, and miscellaneous charges.[39] The declines for these specific charges occurred *above and beyond* the average decline in total arrest charges.
- A relatively small but still statistically significant reduction occurred post-March 2020 in serious arrest charges (e.g., violence, weapon, and Part I offenses), including a 14.6% decline in total violence-related charges, a 17.2% decline in weapons-related charges, and a 31.2% decline in Part I charges.[40] Charges for Part I violence increased 12.3%.

[39] Of the arrests involving traffic-related charges: 68.2% were DUI, and 15.2% were moving and/or texting violations. Drug and alcohol charges were comprised primarily of specific drug charges (40% of which were amphetamine charges) and alcohol possession (roughly 8% of all charges in this category). Since DUI charges were the product of a traffic stop, DUI charge counts were counted as traffic charge counts, and not drug/alcohol charge counts (since those charges were primarily comprised of public intoxication, possession, etc.). Other miscellaneous arrests were fewer than four arrests per month (89% for child neglect). Obstruction charges were primarily failure to appear in court (55%), failure to comply with judicial order (22%) and contempt of court (9%).

[40] The percentages were calculated as follows: violent percentage change = (133.0-155.7)/155.7 = -14.6%.The same formula was for all other charge-specific arrest categories.

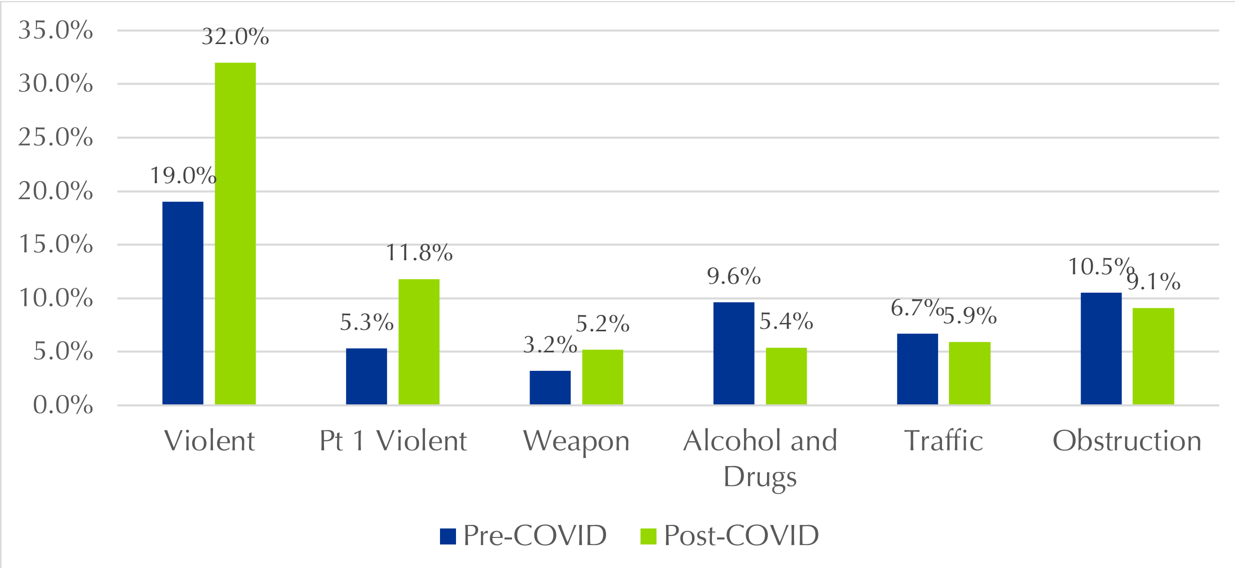
Thus, despite a consistent, pre-existing upward trend in reported criminal offenses, the data shows a significant reduction in arrests post-March 2020, with roughly half of the decline observed for more discretionary charges (e.g., obstruction of justice, drugs and alcohol, and nuisance offenses), and smaller reductions in more serious charges (e.g., violence and weapon-related charges).

Figure 3.10. Pre & Post-COVID Bivariate Change in Arrest Charge Categories



An alternative presentation of these changes is shown in **Figure 3.11** below, where the overall percentage for each arrest category is displayed by comparing the two time periods. For example, 19% of all arrests in the pre-COVID period were for violent offenses, compared to 32% in the post-COVID period. Conversely, drug and alcohol arrests comprised nearly 10% of all arrests in the pre-COVID period but only 5.4% in the post-COVID period. In short, although overall APD arrests decreased, the proportion of those arrests post-COVID for more serious offenses increased.

Figure 3.11. Pre & Post-COVID Bivariate Change in the Proportion of Total Arrests by Charge Category



APD USE OF FORCE, JAN 2017–DEC 2022

As noted in Section 2, use of force can be counted at multiple units of analysis. The NPI team aggregated information to the individual level from data tables at the incident, officer, subject, and force action levels. Like the arrest data, a single person could have force used against them more than once over the six-year period. In these situations, each incident involving the same person is counted as an independent individual experiencing use of force. In addition, multiple individuals could be involved in a single incident (i.e., on the same date/time at the same location, involving the same circumstances). In these situations, each subject is counted as an independent individual experiencing use of force.

Using this definition, across the APD, officers used force against 3,783 individuals during the six-year study period. Of these, 2,608 individuals who had force used against them were arrested (68.9%). Roughly 30% of individuals who had force used against them were not ultimately arrested.

APD uses a tiered system to classify types of force. A full description of the tiers of force is included in Section 2, but a brief overview is provided below:

- Tier Zero: Statutory Use or Display of Force (including pointing of a firearm or pointing of a less lethal weapon or OC)
- Tier One: Use of Force with No or Minor Injury/ Use of Restraint

- Tier Two: Use of weapon other than a deadly weapon to overcome resistance or when subject is injured by member's use of force requiring professional medical treatment
- Tier Three: Use of a deadly weapon, deadly force, or potentially deadly force regardless of any injury

The NPI team examined only Tiers 1–3 use of force reports since Tier Zero does not result in a force report. Thus, the current report cannot assess the frequency, patterns, or trends associated with the APD's use of pointing a firearm.[41] **Figure 3.12** displays the overall distribution of use of force by tier.[42] For the six-year study period, most individuals had less severe force used against them, with 73.1% categorized as Tier One, 25.9% as Tier Two, and only 1.0% as the most serious (Tier Three).

Figure 3.12. Highest Tier Force Experienced by Individuals Who Had Force Used Against Them, 2017–2022 (n = 3,783)

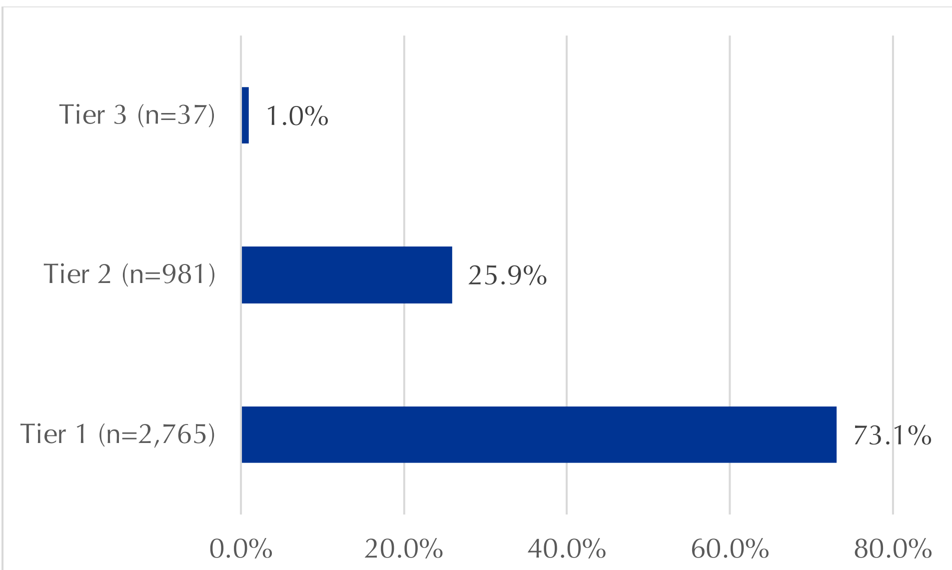


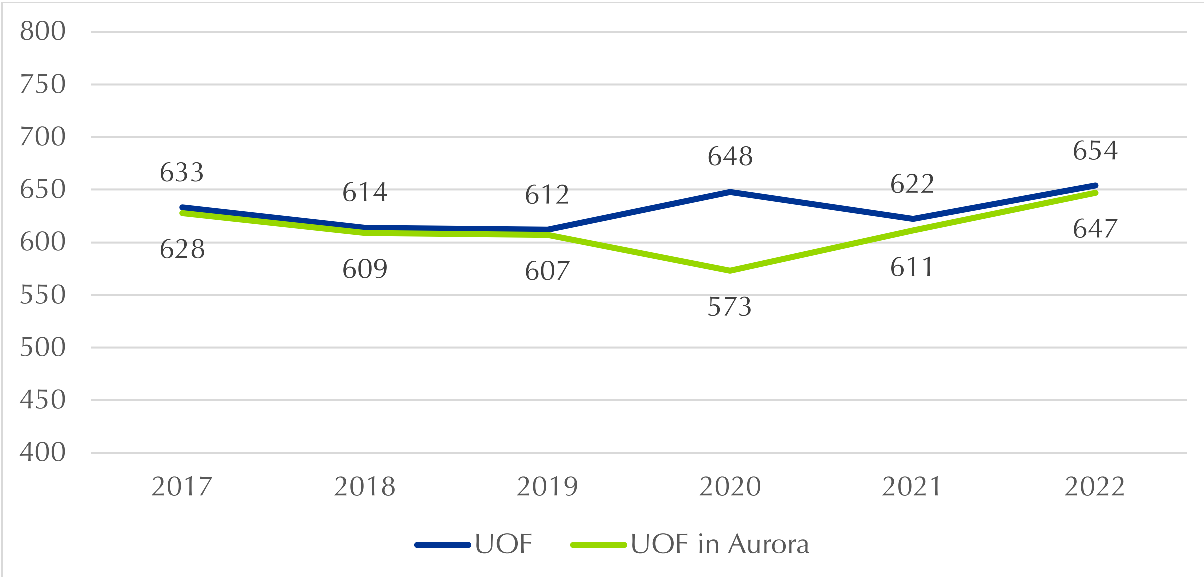
Figure 3.13 shows the distribution of annual counts across the six-year period of the 3,783 individuals who had force used against them by the APD. The annual number of subjects who had force used against them by APD officers remained relatively stable

[41] Research in another Colorado police department showed pointing of a firearm comprised approximately two-thirds of the department's use of force and varied across racial/ethnic groups (Brown et al., 2022).

[42] Due to issues linking subject and incident data (described in Section 2), the NPI team relied on the tier level reported at the incident level. Therefore, the percentages in Figure 3.12 represents the *highest tier* for the overall incident; this may not necessarily apply for the 10% of individuals involved in multi-subject use of force incidents.

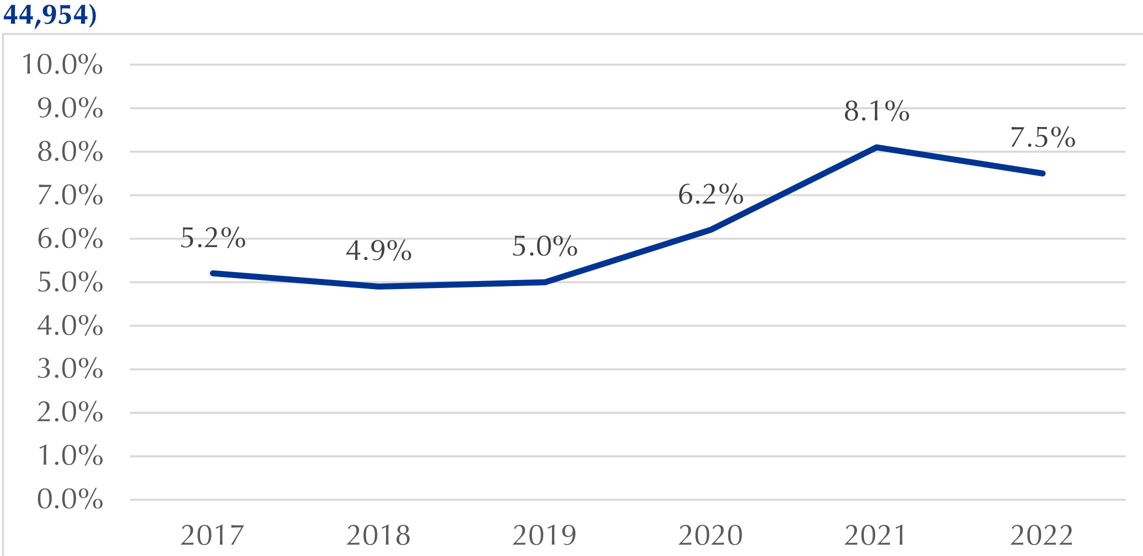
(each year's use of force count is within 5% of the prior year's count). Note in 2020, 75 individuals who had force used against them by APD officers were involved in incidents outside the City of Aurora.

Figure 3.13. Annual Counts of Individuals Who Experienced Use of Force (n = 3,783)



As shown in Figure 3.13 above, in contrast to the trends reported for criminal summonses and arrests, the number of individuals who had force used against them has not significantly decreased. Rather, given the significant reductions in the number of arrestees, the percentage of arrestees who have force used against them significantly *increased* over time (see **Figure 3.14**).

Figure 3.14. Annual Percentages of Arrested Individuals Who Experienced Use of Force (n = 44,954)



To further explore the trends in APD use of force, **Table 3.6** below reports the number of individuals who had force used against them by APD officers within different organizational boundaries.[43] As noted above, in 2020, 75 individuals who had force used against them by APD officers were involved in incidents outside the City of Aurora; 65 of these occurred in Denver. Most of these were related to APD's participation in the police response to protests following the death of George Floyd. Where appropriate, these uses of force are removed from analyses (and noted in the text).

Table 3.6. Annual Counts of Individuals Who Experienced Use of Force by Year (n = 3,783), 2017–2022

	Overall	2017	2018	2019	2020	2021	2022
APD	3,783	633	614	612	648	622	654
District 1	1,543	298	280	250	198	255	262
District 2	1,200	195	196	207	195	188	219
District 3	506	81	76	90	81	85	93
Outside of City	108	5	5	5	75	11	7
Unknown	426	54	57	60	99	83	73

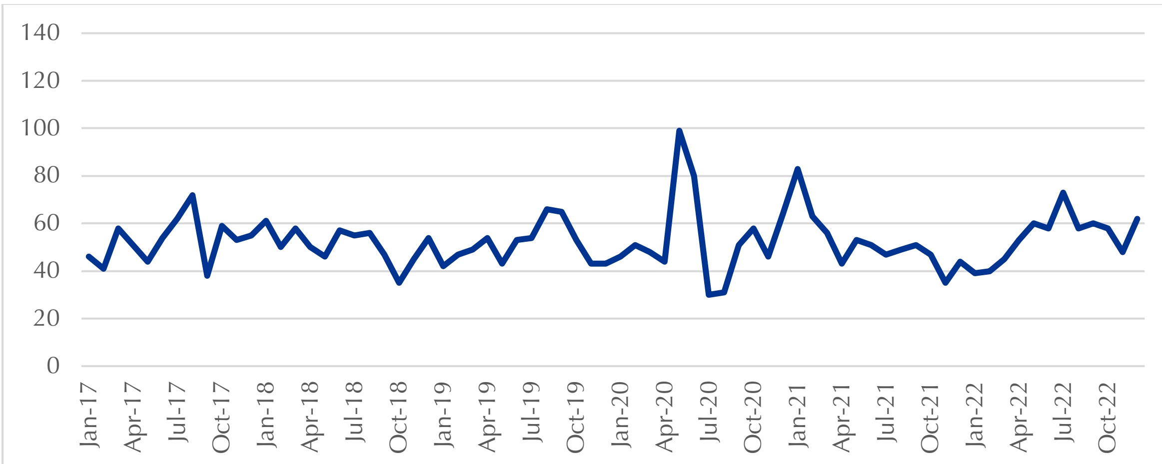
As with criminal summonses and arrests, the NPI team conducted interrupted time series analyses on the monthly counts of individuals who experienced use of force. **Figure 3.15** shows these monthly counts. Although there is seasonal fluctuation and a one-month spike in the number of individuals experiencing use of force, at the bivariate level, the averages pre- and post-COVID and pre- and post-George Floyd protests are similar.[44] Similarly, the time series analysis demonstrated no significant disruption corresponding to any of the examined seminal events (results available upon request). In sum, unlike the significant shifts in the time series for criminal summonses and arrests, neither the

[43] The APD is organized into three patrol districts. The NPI team had intended to examine outcomes of interest at the district level whenever possible to illustrate similarities and differences to assist APD administrators in identifying outliers. Unfortunately, this was not a viable option for two reasons. First, as noted in Section 2, 11.3% of all individuals who had force used against them had missing or incomplete address information in the use of force data, and the district location of their use of force was not able to be determined. Second, there is limited variability across the three districts due to their large spatial distributions. An alternative approach is to use smaller units of analysis; however, in Aurora, the 27 subarea zones could not be used due to limited statistical power.

[44] The average number of individuals who experienced force pre-COVID was 51.4 compared to 53.9 post-COVID. The average number of individuals who experienced force pre-Floyd protests was 51.2 compared to 54.2 post-Floyd protests.

onset of the COVID-19 pandemic nor any other seminal event resulted in a significant, sustained shift in the number of individuals who had force used against them.

Figure 3.15. Monthly Counts of All Individuals Who Had Force Used Against Them, 2017–2022 (n = 3,783)



SECTION SUMMARY

In this section, the NPI team examined trends over time for criminal incidents, criminal summonses, arrests, and use of force using descriptive, bivariate, and interrupted time series analyses. The following key findings are noted.

(1) Crime, especially serious crime (Part I offenses) and serious violent crime (Part I violent offenses), significantly increased in Aurora from 2017 to 2022.
(2) APD officers issued 20,922 criminal summonses from 2017 to 2022. Criminal summonses declined linearly from 2017 to 2022 but were also significantly reduced by 11.2% by the onset of the COVID-19 pandemic (over and above the pre-established linear decline).
(3) APD officers arrested 44,954 individuals from 2017 to 2022. The onset of the COVID-19 pandemic significantly disrupted APD arrest activity, reducing it by approximately 50%. The reduction was primarily driven by decreases in arrests for less serious offenses. The overall proportion of arrests post-COVID increased for more serious and violent offenses.
(4) APD officers used force against 3,783 individuals from 2017 to 2022. The annual number of subjects who had force used against them by APD remained relatively stable, but the percentage of arrestees who experienced use of force significantly increased because of the decline in the number of arrests. Unlike criminal summonses and arrests, use of force was not significantly disrupted by the onset of the COVID-19 pandemic.

SECTION 4: ARREST & USE OF FORCE BENCHMARK COMPARISONS

This section examines the racial/ethnic composition of the population of arrested individuals and those who had force used against them by APD officers. As described in Section 2, understanding whether racial/ethnic disparities exist in police enforcement outcomes requires comparing the percentages of individuals with those outcomes to a valid benchmark group. A benchmark should estimate similarly situated people at risk of experiencing these outcomes, assuming no bias exists (Engel & Calnon, 2004; PERF, 2021; Tillyer et al., 2010). A benchmark analysis involves comparing the percentage of racial and ethnic groups who experience arrests or force and the percentage of racial and ethnic groups in the estimated population of similarly situated people. Section 2 summarized the strengths and limitations of various benchmarks to approximate those at risk of experiencing these outcomes, including the calculation of disproportionality indices and disparity ratios for interpreting benchmark comparisons.

This section provides arrest disparity ratios based on four benchmarks and use of force disparity ratios based on eight benchmarks. Given the differences in reported crimes, criminal summonses, arrests, and uses of force across periods identified using interrupted time series analyses in Section 3, the NPI team also calculated and compared disparity ratios for two distinct periods:

- Period 1: Jan 2019–Mar 2020 (Pre-COVID)
- Period 2: Apr 2020–Dec 2022 (Post-COVID)

Based on the known limitations of various benchmarks (see Section 2), the NPI team relied on several benchmarks to provide a more holistic picture of racial/ethnic disparities across different data sources.

ARREST BENCHMARKS

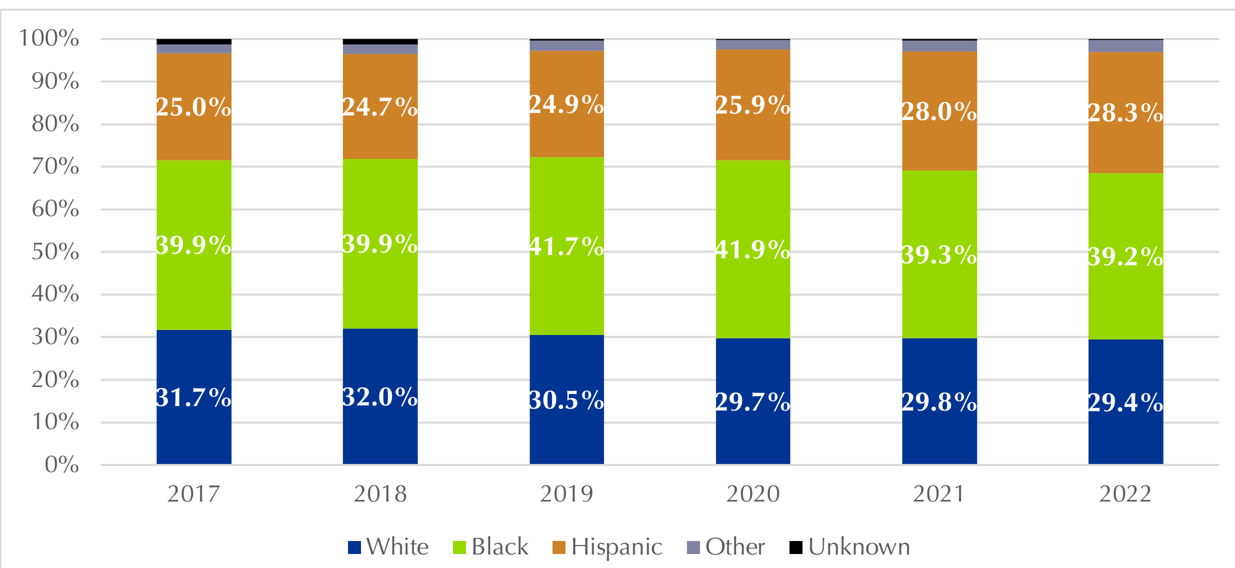
Table 4.1 shows the percentage of arrested individuals by race/ethnicity in the study period. Of those arrested, 40.3% were Black, 30.8% were White, and 25.8% were Hispanic. Other or unknown race/ethnicity categories comprised the remaining 3.1%.

Table 4.1. Race/Ethnicity of Individuals Arrested by Year, 2017–2022 (n = 3,783)

Arrest	Race/Ethnicity N (%)				
	White	Black	Hispanic	Other	Unknown
Total 2017-2022 (n = 44,954)	13,838 (30.8%)	18,137 (40.3%)	11,579 (25.8%)	1,069 (2.4%)	331 (0.7%)
2017 (n = 9,780)	3,098 (31.7%)	3,906 (39.9%)	2,445 (25.0%)	202 (2.1%)	129 (1.3%)
2018 (n = 10,277)	3,289 (32.0%)	4,098 (39.9%	2,534 (24.7%)	221 (2.2%)	135 (1.3%)
2019 (n = 9,280)	2,826 (30.5%)	3,874 (41.7%)	2,313 (24.9%)	236 (2.5%)	31 (0.3%)
2020 (n = 4,951)	1,470 (29.7%)	2,074 (41.9%)	1,282 (25.9%)	116 (2.3%)	9 (0.2%)
2021 (n = 4,689)	1,398 (29.8%)	1,842 (39.3%)	1,313 (28.0%)	120 (2.6%)	16 (0.3%)
2022 (n = 5,977)	1,757 (29.4%)	2,343 (39.2%)	1,692 (28.3%)	174 (2.9%)	11 (0.2%)

Figure 4.1 graphically displays the percentage of arrested individuals by race/ethnicity by year. As shown, across all years, Black individuals represented the largest percentage of APD arrests, while White individuals consistently comprised the second highest percentage of arrested individuals, followed by Hispanic individuals. The racial/ethnic distribution of arrested individuals was relatively consistent over time, although the percentage of Hispanic arrestees slightly increased in 2021 and 2022. Individuals of other or unknown races represented a small percentage of arrestees across all years.

Figure 4.1. Race/Ethnicity of Arrested Individuals by Year, 2017–2022 (n = 3,783)



As described in Section 2, simply knowing the racial/ethnic breakdown of arrested individuals is not useful without a comparison to a valid benchmark. **Table 4.2** displays the values of the disproportionality indices and disparity ratios comparing the percentage of racial and ethnic groups among arrestees with the percentage of racial and ethnic groups among four comparison data sources (or benchmarks)[45] described in Section 2. These include:

(1) residential population
(2) all crime suspects
(3) crime suspects of Part I offenses
(4) crime suspects of Part I violent offenses

Table 4.2. Comparison of APD Arrest Racial/Ethnic Disparity Ratios Across Benchmarks

	Percent Race/Ethnicity			Disproportionality Indices			Disparity Ratios	
	White	Black	Hispanic	White	Black	Hispanic	Black	Hispanic
% Arrests (N = 44,954)[46]	30.8% (13,838)	40.3% (18,137)	25.8% (11,579)	--	--	--	--	--
Benchmark 1: % Residential Population	43.5%	16.6%	29.0%	0.71	2.43	0.89	**3.43**	**1.26**
Benchmark 2: % Suspect Population (All Crimes)	34.6%	36.5%	21.9%	0.89	1.10	1.18	**1.24**	**1.33**
Benchmark 3: % Suspect Population (Part I Crimes)	31.6%	39.0%	22.0%	0.98	1.03	1.17	**1.06**	**1.20**
Benchmark 4: % Suspect Population (Part I Violent Crime)	24.5%	45.5%	24.4%	1.26	0.89	1.06	**0.70**	**0.84**

To aid in comparing across benchmarks, **Figure 4.2** visually displays the arrest disparity ratios for Black and Hispanic individuals based on the four benchmarks reported in Table 4.2. The red line indicates no racial/ethnic disparities detected (DR = 1.0). Bars

[45] Unlike use of force benchmark comparisons presented later in this section, arrest-based benchmarks are not included in Table 4.2 or Figure 4.2 since the numerator is the racial/ethnic percentages of all arrests.
[46] Not displayed in tabular or graphic format are 1,400 arrested individuals who were reported to belong to "other" racial/ethnic groups or were of unknown race/ethnicity.

above the 1.0 threshold show that Black and Hispanic individuals were *more likely* than White individuals to be arrested (based on the respective benchmark). In contrast, bars under the red line demonstrate that Black and Hispanic individuals were *less likely* than White individuals to be arrested (based on the respective benchmark).

Figure 4.2. Comparison of APD Arrest Racial/Ethnic Disparity Ratios Across Benchmarks

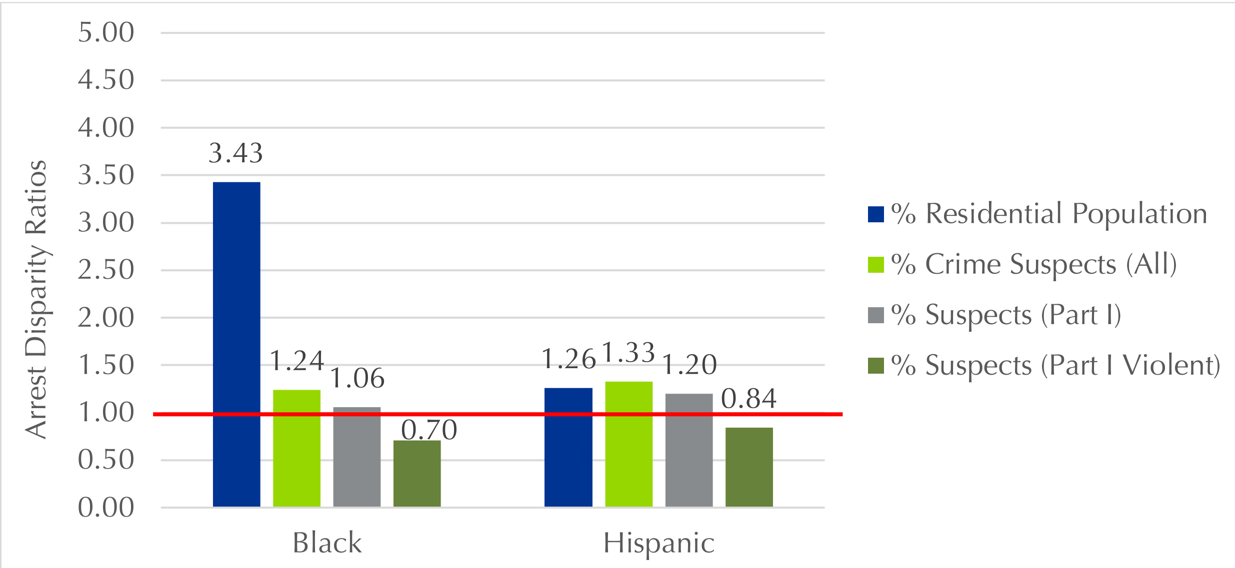


Table 4.2 and Figure 4.2 show that the highest disparity ratio for Black individuals (3.43 =2.43/0.71) results from census-based residential population comparisons. Black individuals were 3.43 times more likely to be arrested than White individuals based on each group's representation in the residential population. A similar finding, though smaller in magnitude, is noted for Hispanic individuals. The disparity ratio for Hispanic individuals was 1.26, so Hispanic individuals were slightly more likely to be arrested compared to White individuals based on residential population statistics.

In Table 4.2, the NPI team also presents three comparisons of arrested individuals to those reported as criminal suspects (all suspects, Part I crime suspects, Part I violent crime suspects). Using all crime suspects as the benchmark, the disparity ratio for Black individuals is 1.24, indicating that Black individuals were somewhat more likely than White individuals to be arrested. The disparity ratios for Black individuals are closer to 1.0 when the criminal suspect benchmark is limited to Part I criminal suspects (DR=1.06) and less than 1.0 when based on Part I violent crime suspects (DR=0.70). This highlights that Black individuals were *less likely* than White individuals to be arrested based on their groups' representation among the violent criminal suspect population. For Hispanic individuals, the disparity ratio based on all crime suspects is 1.33, and all Part I suspects is 1.20, indicating that Hispanic individuals were somewhat more likely than White individuals to be arrested based on each group's representation among the suspect comparison sources. Compared to the Part I violent crime suspect benchmark, the disparity ratio for Hispanics is less than 1.0, indicating they are less likely than White individuals to be arrested.

The arrest disparity ratios demonstrate that comparing residential population-based benchmarks produces a vastly different picture of racial/ethnic disparities in arrests than suspect-based benchmarks that better estimate individuals *at risk* of interacting with and being arrested by the police. The validity of using ccensus-based benchmarks has been routinely called into question by policing scholars (Alpert et al., 2004; Engel & Calnon, 2004; Engel et al., 2023; Fridell, 2004; Geller et al., 2021; Smith et al., 2019).

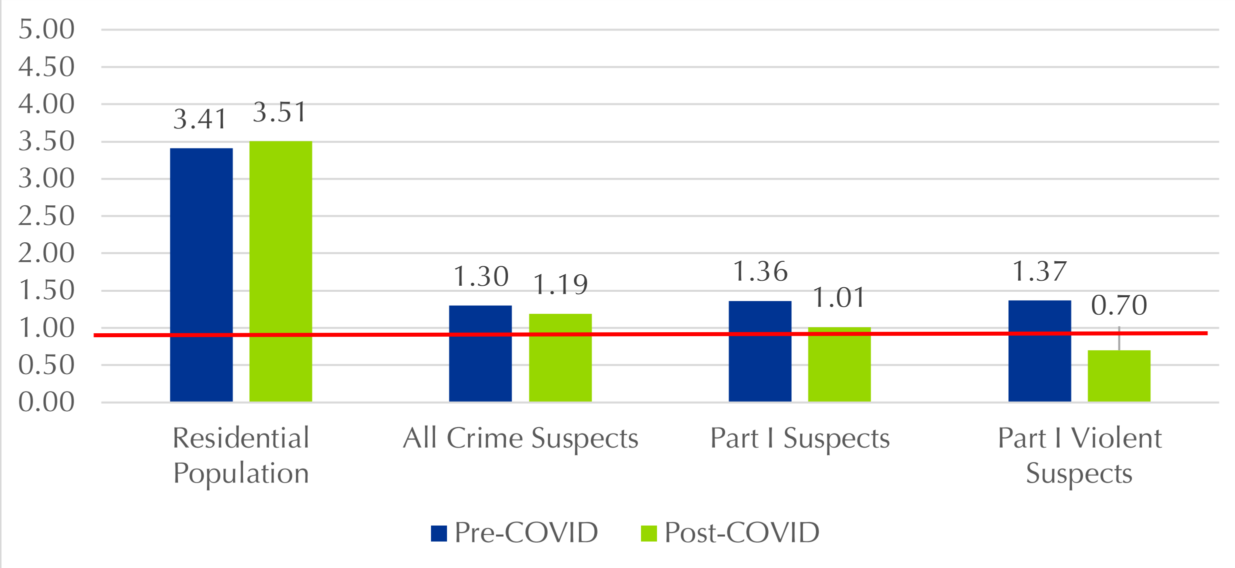
Arrest Benchmark Comparisons Over Time

The interrupted time series analyses presented in Section 3 demonstrated significant shifts in enforcement activities after the onset of the COVID-19 pandemic. None of the other seminal events the NPI team examined significantly shifted enforcement activities. Given the clear differences in enforcement activities pre- and post-COVID, separating these periods for additional analyses is helpful. Specifically, the NPI team calculated and compared disparity ratios for two distinct time periods:

- Period 1: Jan 2019–Mar 2020 (Pre-COVID)
- Period 2: Apr 2020–Dec 2022 (Post-COVID)

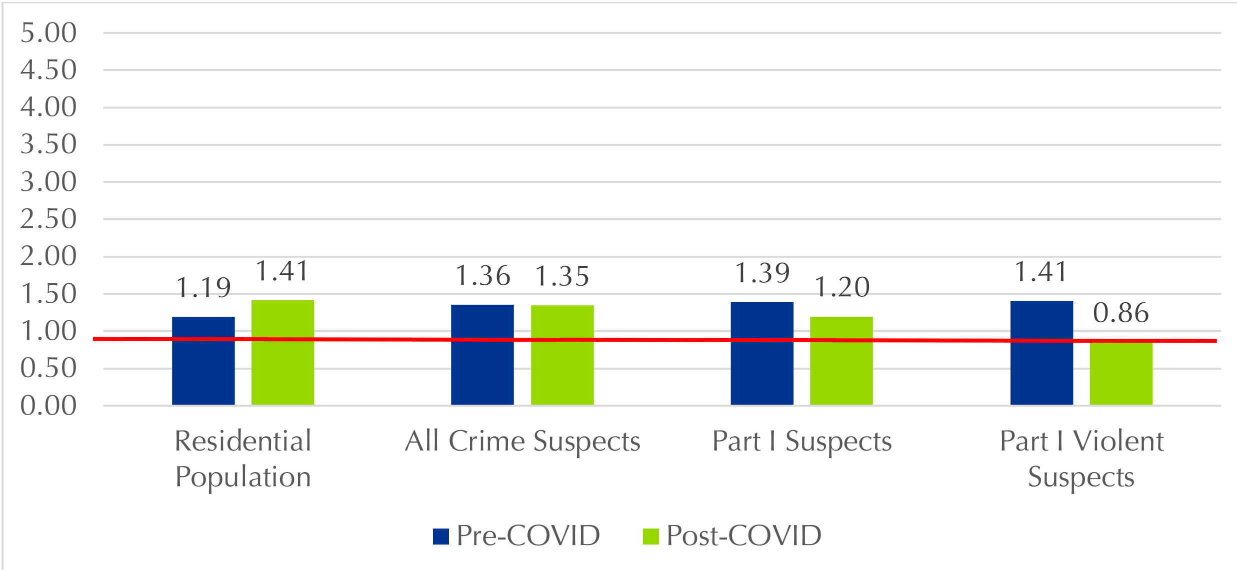
Figure 4.3 displays the arrest disparity ratios for Blacks compared to Whites, while **Figure 4.4** displays the same information for Hispanics compared to Whites. The table documenting these calculations is included in the Appendix. As shown, the disparity ratio based on residential population data from the census slightly increased after the onset of COVID-19 from 3.41 to 3.51. However, across all suspect-based benchmarks, the arrest disparity ratios for Blacks compared to Whites are *lower* for the post-COVID period. The post-COVID disparity ratio based on Part I suspects suggests that Blacks are equally likely to be arrested compared to Whites based on Part I suspects and less likely to be arrested based on Part I violent suspects. This suggests that the decline and sustained reduction in arrests documented in Section 3 has also reduced arrest disparity ratios for Blacks.

Figure 4.3. APD Arrest Disparity Ratios Comparing Blacks to Whites, Pre-COVID vs. Post-COVID



For Hispanic individuals, the arrest disparity ratios show a similar pattern in the census-based benchmark, which increases from 1.19 to 1.41 after the onset of COVID-19. There is virtually no change in the disparity ratio based on all crime suspects; Hispanics remain approximately 1.35 times more likely to be arrested than Whites compared to their representation in the crime suspect population. The disparity ratio based on Part I suspects is slightly reduced after the onset of COVID-19 from 1.39 to 1.20. Finally, similar to the pattern for Blacks, the post-COVID disparity ratio based on Part I violent suspects shows that Hispanics were less likely to be arrested than Whites.

Figure 4.4. APD Arrest Disparity Ratios Comparing Hispanics to Whites, Pre-COVID vs. Post-COVID



USE OF FORCE BENCHMARKS

Table 4.3 shows the percentage of individuals who had force used against them by race/ethnicity in the six-year period. Of those who had force used against them, 43.1% were Black, 33.5% were White, and 15.3% were Hispanic. Other or unknown race/ethnic categories comprised the remaining 8.2%.

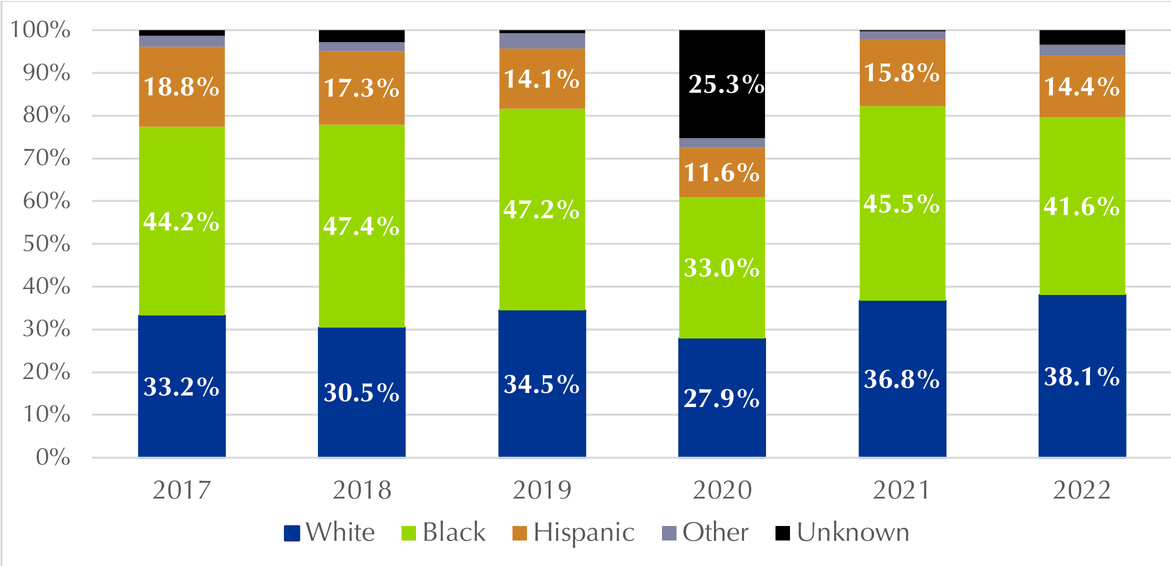
Table 4.3. Race/Ethnicity of Subjects Who Had Force Used Against Them by Year, 2017–2022 (n = 3,783)

Use of Force	Race/Ethnicity N (%)				
	White	Black	Hispanic	Other	Unknown
Total 2017-2022 (n = 3,783)	1,267 (33.5%)	1,629 (43.1%)	578 (15.3%)	93 (2.5%)	216 (5.7%)
2017 (n = 633)	210 (33.2%)	280 (44.2%)	119 (18.8%)	16 (2.5%)	8 (1.3%)
2018 (n = 614)	187 (30.5%)	291 (47.4%)	106 (17.3%)	13 (2.1%)	17 (2.8%)
2019 (n = 612)	211 (34.5%)	289 (47.2%)	86 (14.1%)	22 (3.6%)	4 (0.7%)
2020 (n = 648)	181 (27.9%)	214 (33.0%)	75 (11.6%)	14 (2.2%)	164 (25.3%)
2021 (n = 622)	229 (36.8%)	283 (45.5%)	98 (15.8%)	11 (1.8%)	1 (0.2%)
2022 (n = 654)	249 (38.1%)	272 (41.6%)	94 (14.4%)	17 (2.6%)	22 (3.4%)

NOTE: Other race includes Asian, Hawaiian or Pacific Islander, and American Indian or Alaska Native

Figure 4.5 graphically displays the percentage of individuals who had force used against them by race/ethnicity and year. As shown, across all years, Black individuals represented the largest percentage of those who had force used against them by APD officers. Whites consistently comprised the second-highest percentage of individuals who had force used against them. Hispanics were the third most common racial/ethnic group represented among those who had force used against them except in 2020. The racial/ethnic distribution of individuals who had force used against them was relatively consistent, except in 2020 when 25.3% of individuals were reported to be of unknown race, and the percentages of all other race/ethnicity categories decreased.

Figure 4.5. Race/Ethnicity of Subjects With Force Used by Year, 2017–2022 (n = 3,783)



As described in Section 2, simply knowing the racial/ethnic breakdown of individuals who had force used against them is not useful without a comparison to a valid benchmark. **Table 4.4** includes disproportionality indices and disparity ratios that compare the percentage of racial and ethnic groups who experienced force with the total percentage of racial/ethnic groups within eight comparison data sources (or benchmarks) previously described in Section 2. These include:

(1) residential population
(2) criminal summonses
(3) all arrestees
(4) arrestees for Part I offenses
(5) arrestees for Part I violent offenses
(6) all crime suspects
(7) crime suspects of Part I offenses
(8) crime suspects of Part I violent offenses

Table 4.4. Comparison of APD Use of Force Racial/Ethnic Disparity Ratios Across Benchmarks

	Percent Race/Ethnicity			Disproportionality Indices			Disparity Ratios	
	White	Black	Hispanic	White	Black	Hispanic	Black	Hispanic
% Use of Force (n = 3,783)[47]	33.5% (1,267)	43.1% (1,629)	15.3% (578)	--	--	--	--	--
Benchmark 1: % Residential Population	43.5%	16.6%	29.0%	0.77	2.60	0.53	3.37	0.69
Benchmark 2: Criminal Summons Population	33.4%	36.3%	25.5%	1.00	1.19	0.60	1.18	0.60
Benchmark 3: % Arrestee Population (All crimes)	30.8%	40.3%	25.8%	1.09	1.07	0.59	0.98	0.55
Benchmark 4: % Arrestee Population (Part I Crimes)	29.6%	43.2%	24.3%	1.13	1.00	0.63	0.88	0.56
Benchmark 5: % Arrestee Population (Part I Violent Crimes)	24.7%	46.6%	25.8%	1.36	0.92	0.59	0.68	0.44
Benchmark 6: % Suspect Population (All Crimes)	34.6%	36.5%	21.9%	0.97	1.18	0.70	1.22	0.72
Benchmark 7: % Suspect Population (Part I Crimes)	31.6%	39.0%	22.0%	1.06	1.10	0.69	1.04	0.65
Benchmark 8: % Suspect Population (Part I Violent Crimes)	24.5%	45.5%	24.4%	1.37	0.95	0.63	0.69	0.46

To aid in comparing across benchmarks, **Figure 4.6** displays the use of force disparity ratios for Black and Hispanic individuals based on each of the eight benchmarks reported in Table 4.4. Again, the red line indicates no racial/ethnic disparities detected (DR = 1.0). Bars above the 1.0 threshold show that Black and Hispanic individuals were *more likely* than White individuals to have force used against them (based on the respective benchmark), while bars under the red line demonstrate that Black and

[47] Not displayed in tabular or graphic format are 309 individuals of "other" or unknown race/ethnicity who had force used against them.

Hispanic individuals were *less likely* than White individuals to have force used against them (based on the respective benchmark).

Figure 4.6. Comparison of APD Use of Force Racial/Ethnic Disparity Ratios Across Benchmarks

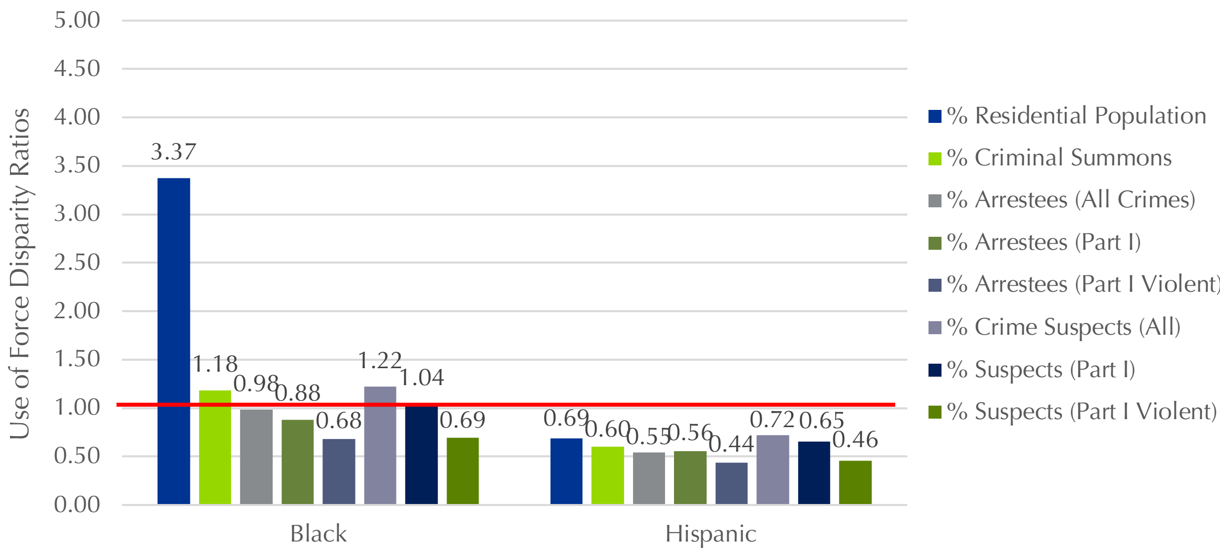


Table 4.4 and Figure 4.6 show that the highest disparity ratio for Black individuals (3.37=2.60/0.77) results from census-based residential population comparisons. Black individuals were 3.37 times more likely to be arrested than White individuals based on each group's representation in the residential population. By contrast, the disparity ratio for Hispanic individuals was 0.69. Thus, Hispanic individuals were less likely to have force used against them compared to White individuals based on the underlying residential population. When the residential population is used as a benchmark comparison to estimate risk for police use of force, Black individuals, but not Hispanic individuals, were overrepresented in use of force compared to their White counterparts.

Comparing individuals who had force used against them to those who received a criminal summons shows a disparity ratio for Black individuals that is much closer to 1.0 (DR=1.18), while for Hispanic individuals, the criminal summons-based disparity ratio is less than 1.0 (DR=0.60). Next, the NPI team examined use of force by race/ethnicity compared to the race/ethnicity of the APD arrestee population from 2017 to 2022. As shown in Table 4.4 and Figure 4.6, the disparity ratio based on total arrests is 0.98 for Black individuals and 0.55 for Hispanic individuals. In comparison to the residential population-based disparity ratios, the summons- and arrest-based benchmark comparisons illustrate that using benchmarks that estimate individuals *at risk* of police use of force produces a different picture of racial/ethnic disparities in use of force.

When the benchmark is changed to only Part I crime arrests or Part I violent crime arrests, the disparity ratios drop even further for Blacks to 0.88 and 0.68, respectively.

Disparity ratios less than 1.0 indicate that Black and Hispanic individuals were underrepresented among individuals who had force used against them compared to White individuals based on their representation in the Part I crime and Part I violent crime arrestee populations.

Finally, the NPI team conducted benchmark analyses based on criminal suspect data, using the race/ethnicity recorded by APD for individuals reported as criminal suspects by the public when reporting criminal events. Using all crime suspects as the benchmark, the disparity ratio for Blacks is 1.22, indicating that Black individuals were somewhat more likely than White individuals to experience force. The disparity ratios for Black individuals are closer to 1.0 when the criminal suspects benchmark is limited to Part I criminal suspects (DR=1.04) and less than 1.0 when based on Part I violent crime suspects (DR=0.69), showing that Black individuals were *less likely* than White individuals to have force used against them based on their groups' representation among the violent criminal suspect population. For Hispanic individuals, regardless of which suspect benchmark is used, the disparity ratios are all less than 1.0, indicating they are *less likely* than White individuals to experience force based on their representation in the suspect-based benchmarks.

These findings, particularly for Blacks who experienced use of force by APD officers, are consistent with previous studies that have compared variation in racial/ethnic disparities across different benchmarks. As illustrated, the use of force disparity ratios created using non-census data sources are all close to, or less than 1.0, indicating that there is limited or no disparity between Black or Hispanic individuals' likelihood of having force used against them in comparison to White individuals. Additionally, using some benchmarks, Black and Hispanic individuals were underrepresented in the use of force population compared to White individuals, given their representation in several benchmark populations.

Use of Force Benchmark Comparisons Over Time

Given the significant shifts in enforcement activities described in Section 3, the NPI team also calculated and compared disparity ratios for two distinct periods:

- Period 1: Jan 2019–Mar 2020 (Pre-COVID)
- Period 2: Apr 2020–Dec 2022 (Post-COVID)

Figure 4.7 displays the use of force disparity ratios for Black individuals compared to White individuals, while **Figure 4.8** displays the same information for Hispanic individuals compared to White individuals. The table documenting these calculations is included in the Appendix. As shown, across all benchmarks, the disparity ratios for Black compared to White individuals are *lower* for the post-COVID period. During Period 2, two of the eight benchmarks are between 1.0 and 1.1, while five are less than 1.0, indicating that Black individuals were *less likely* than White individuals to experience force based on their representation in those comparison populations. This suggests that the decline and sustained reduction in arrests documented in Section 3 has also reduced use of force disparity ratios for Black individuals.

For Hispanic individuals, the use of force disparity ratios were all 1.0 or less in both the pre-COVID and post-COVID periods. This is consistent across all benchmarks. Again, this indicates that Hispanic individuals experience less use of force than White individuals, given the expected rate of force based on each group's representation among Aurora residents, those who received criminal summonses, arrestees, or reported criminal suspects.

Figure 4.7. APD Use of Force Disparity Ratios Comparing Blacks to Whites, Pre-COVID vs. Post-COVID

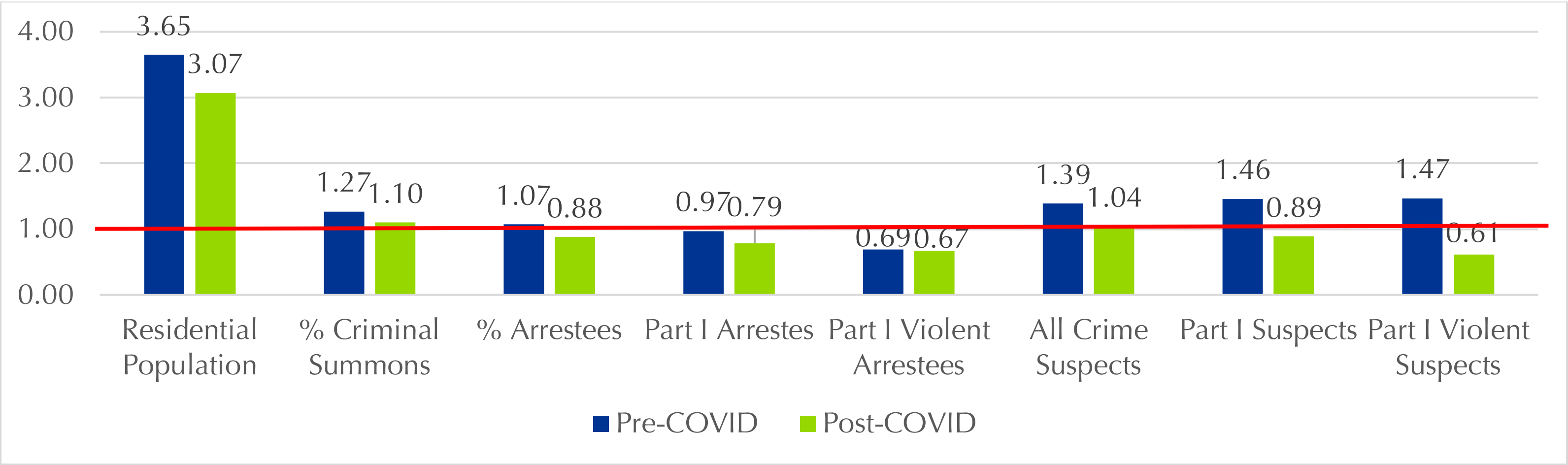
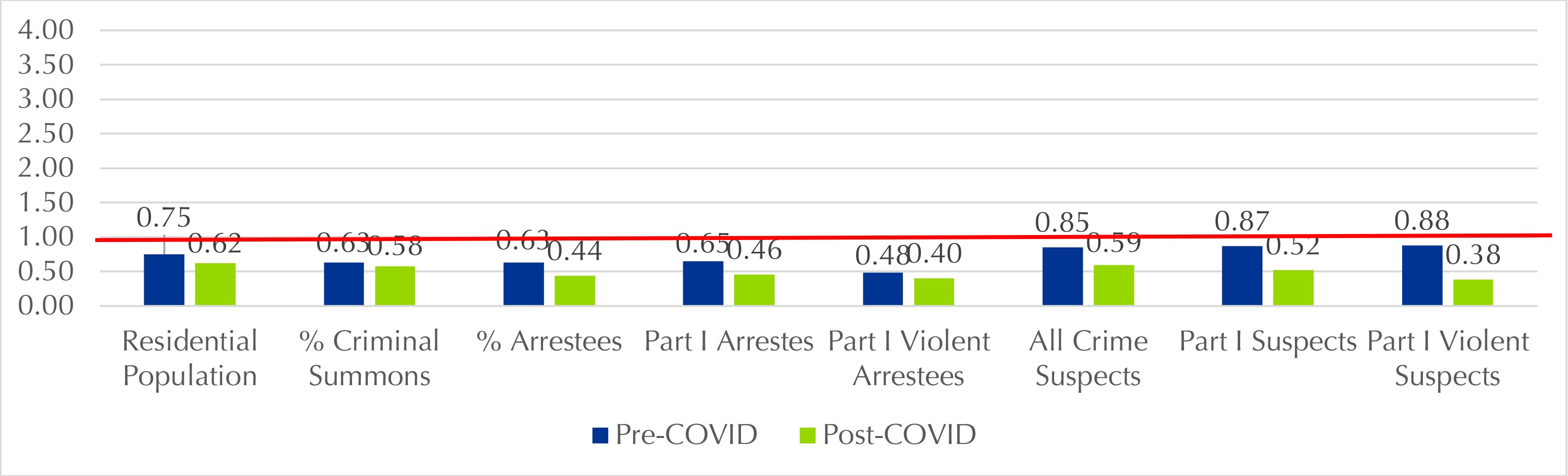


Figure 4.8. APD Use of Force Disparity Ratios Comparing Hispanics to Whites, Pre-COVID vs. Post-COVID



SECTION SUMMARY

All benchmarks have limitations and vary in the extent to which they accurately estimate the population of similarly situated individuals "at risk" of police enforcement actions, assuming no bias exists (Alpert et al., 2004; Engel & Calnon, 2004; Engel et al., 2023; Fridell, 2004; Geller et al., 2021; PERF, 2021; Tillyer et al., 2010). For example, residential population-based benchmarks do not include measures of factors that influence an individual's risk of police enforcement activity, including subjects' resistance, presence of a weapon, and criminal behavior. Similarly, using arrest data as a comparison for use of force benchmark analyses may underestimate disparities because of the possible (unmeasured) bias in who is arrested (Geller et al., 2021; Knox et al., 2020a, 2020b). Arrest data is also challenging as an independent benchmark because arrest disparities are passed on to the next analysis, compounding the possible differences across racial/ethnic groups (Fenton et al., 2020; Knox & Mummolo, 2020).

Based on the known limitations of certain benchmarks, the NPI team relied on various benchmarks using different data sources to provide a more holistic picture of racial/ethnic disparities. The results of the benchmark analyses should be interpreted with caution and consideration of how well each benchmark estimates the "similarly situated" or at-risk population for police enforcement actions. Findings can vary dramatically based on the chosen benchmarks. Previous research suggests that racial/ethnic disparities in use of force are almost always the largest when comparisons are based on residential population and considerably smaller when based on arrest and suspect-based benchmarks that capture the risk of police interactions that may result in use of force (Cesario et al., 2019; Fryer, 2019; Geller et al., 2021; Ross et al., 2020; Smith et al., 2019; Smith et al., 2022; Tregle et al., 2019). Finally, benchmark analyses (regardless of the comparison data source) cannot explain the reasons behind any reported disparities because they do not consider the complex factors that may contribute to differential outcomes across racial and ethnic groups.

For the six-year study period, APD arrested 44,954 individuals. Of these, 40.3% were Black, 30.8% were White, 25.8% were Hispanic, and 3.1% were of other or unknown racial/ethnic backgrounds. The NPI team compared the percentage of racial/ethnic groups' representation in the population of arrested individuals to racial/ethnic groups' representation in four different benchmarks, including (1) residential population, (2) all crime suspects, (3) crime suspects for Part I offenses, and (4) crime suspects for Part I violent offenses. Although the population-based disparity ratios for both Blacks and Hispanics increased after the onset of COVID-19, disparities in arrests for Blacks compared to Whites decreased post-COVID across all suspect-based benchmarks. For Hispanics, two of the three suspect-based benchmarks also declined post-COVID. The post-COVID arrest disparity ratios based on Part I violent suspects show that both Black and Hispanic individuals were less likely to be arrested than their White counterparts.

During the study period, 3,783 individuals had force used against them. Of these, 43.1% were Black, 33.5% were White, 15.3% were Hispanic, 5.7% were of unknown race/ethnicity, and 2.5% were of other racial/ethnic backgrounds. The NPI team compared the percentage of racial/ethnic groups' representation in the population of those who experienced force to racial/ethnic groups' representation in eight different benchmarks, including (1) residential population, (2) criminal summonses, (3) all arrestees, (4) arrestees for Part I offenses, (5) arrestees for Part I violent offenses, (6) all crime suspects, (7) crime suspects for Part I offenses, and (8) crime suspects for Part I violent offenses. Disparities in use of force for Blacks compared to Whites decreased post-COVID across all benchmarks, while no disparities in use of force for Hispanics were evident across all benchmarks either before or after COVID-19 onset.

SECTION 5: PREDICTING USE OF FORCE DURING ARRESTS

Given that use of force is more common during police encounters involving arrests (Garner et al., 2018; Hickman et al., 2008), the overall decline in APD arrests since 2020 is essential for understanding the APD's use of force patterns and trends. In this section, the NPI team compares the percentage of arrestees that have force used against them over time and by racial/ethnic groups. The remainder of this section focuses on using multivariate statistical analyses to better understand what factors predict whether arrested individuals experience use of force.

DESCRIPTIVE ANALYSES

During the six-year study period, APD officers arrested 44,954 individuals during encounters with police. Despite the increased risk for confrontation that these encounters present, most arrestees did not experience use of force by the APD. On average, approximately 5.8% of arrested individuals (n = 2,608) had force used against them. **Table 5.1** below shows the distribution of arrested individuals who had force used against them.

Table 5.1. APD Arrest Counts and Use of Force Counts Within Arrests, 2017–2022 (n = 44,954)

	Overall 2017-2022	2017	2018	2019	2020	2021	2022
Number of Arrestees	44,954	9,780	10,277	9,280	4,951	4,689	5,977
Number of Arrestees with Use of Force	2,608	504	504	463	307	381	449
% Arrestees with Use of Force	5.8%	5.2%	4.9%	5.0%	6.2%	8.1%	7.5%

NOTE: The totals in this table include all arrests, geocoded/mapped or not, and represent 100% of the total distribution of APD arrests.

Note that the individuals included in the statistical analyses in this section only included 68.9% (n = 2,608) of all individuals (n = 3,783) who experienced force during this six-year period. Roughly 30% of the individuals who had force used against them were not arrested by the APD, and therefore, are not included in the arrest database. The analytical techniques used within this section are applied to the arrest database rather than the use of force database. The purpose is to understand what factors predict the likelihood of force in the situations that are most at-risk of involving force. In this case, the population of arrested individuals is known; and we seek to explore what factors predict the likelihood of experiencing force among the 5.8% of arrestees who have force used against them during incidents with police.

Figure 5.1 graphically displays the annual percentage of individuals arrested by APD officers who experienced use of force. It is important to note that although *fewer* individuals had force used against them in 2020–2022 compared to 2017–2019, a *larger* percentage of arrestees had force used against them. As shown in Figure 5.1, the percentage of arrestees who had force used against them was relatively stable between 2017 and 2019 (3-year avg = 5.0%), however, this percentage increased from 2020 to 2022 (3-year avg = 7.3%) as the number of arrests decreased. This represents a percentage change increase of 46% between the two periods (2017–2019 compared to 2020–2022).

Figure 5.1. Percentage of Arrested Individuals Who Experienced Use of Force by Year (n = 44,954 arrestees)

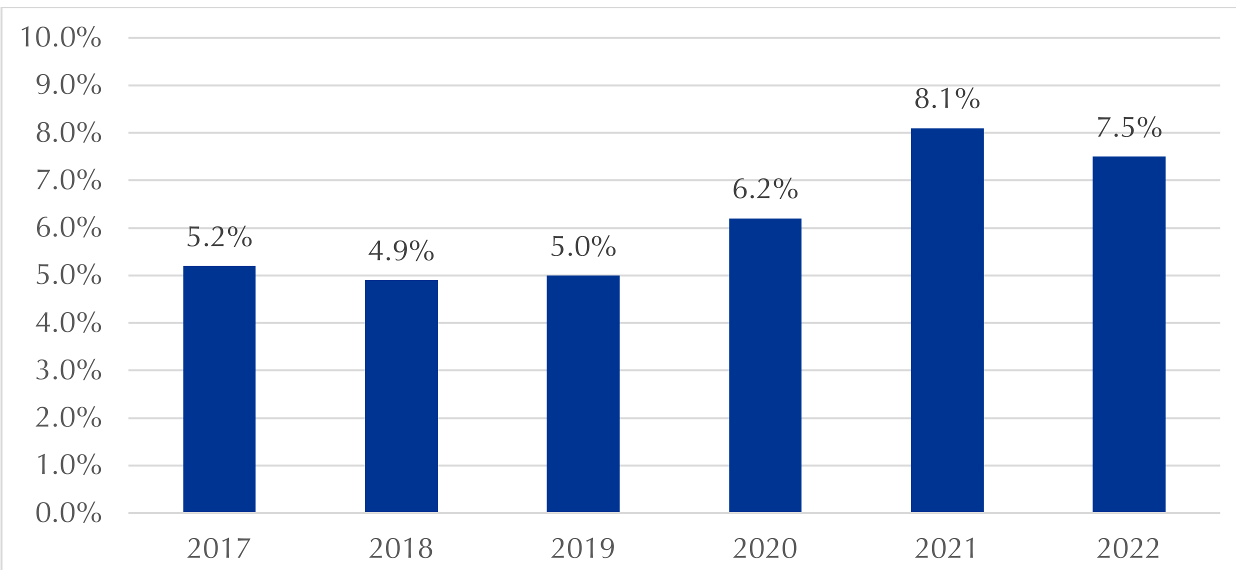


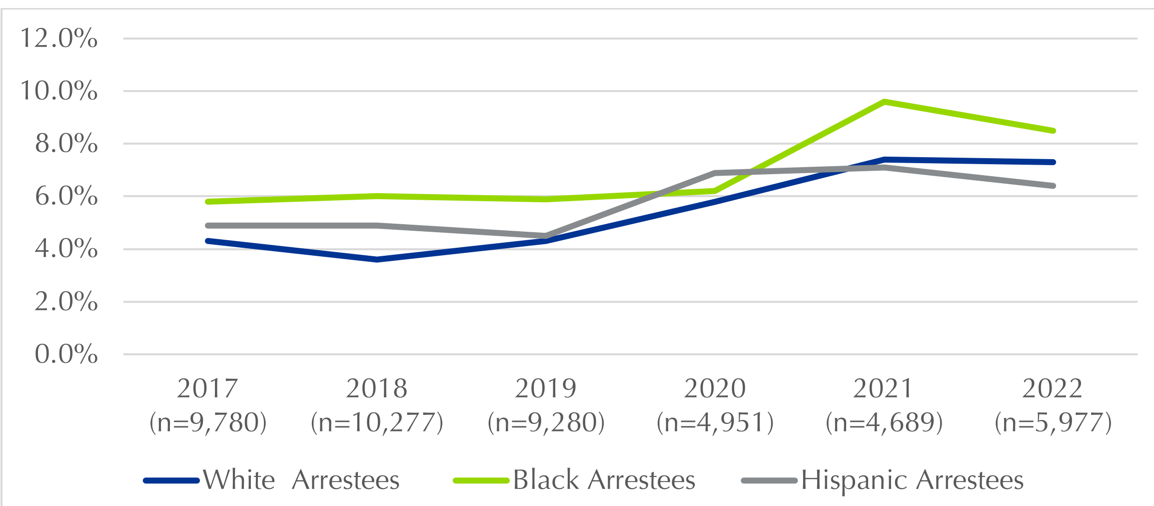
Table 5.2 shows the percentage of arrested individuals who experienced use of force from 2017 to 2022 by race and ethnicity.

Table 5.2. Percentage of Arrested Individuals Who Experienced Use of Force by Race/Ethnicity

	2017-2022 (n = 44,954)	2017 (n = 9,780)	2018 (n = 10,277)	2019 (n = 9,280)	2020 (n = 4,951)	2021 (n = 4,689)	2022 (n = 5,977)
All Arrestees	5.8%	5.2%	4.9%	5.0%	6.2%	8.1%	7.5%
White Arrestees	5.0%	4.3%	3.6%	4.3%	5.8%	7.4%	7.3%
Black Arrestees	6.6%	5.8%	6.0%	5.9%	6.2%	9.6%	8.5%
Hispanic Arrestees	5.5%	4.9%	4.9%	4.5%	6.9%	7.1%	6.4%

As graphically displayed in **Figure 5.3**, the percentage of arrestees who had force used against them increased across all racial/ethnic groups, although there was variation in the magnitude of this increase.

Figure 5.3. Percentage of Arrested Individuals Who Experienced Use of Force by Race/Ethnicity by Year, 2017–2022 (n = 44,954 arrests)



- White Arrestees:
  - Use of force ranges from 3.6% (2018) to 7.4% (2021) of arrestees
  - Average 2017–2019 = 4.1% arrestees with force
  - Average 2020–2022 = 6.8% arrestees with force

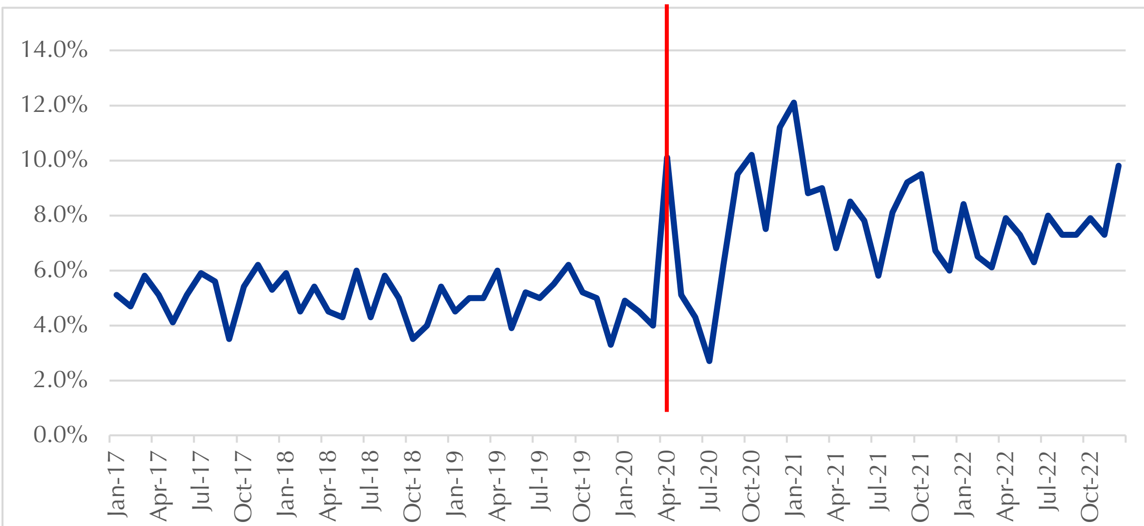
   - Increase of 65.9% arrestees with force (2017–2019 avg vs. 2020–2022 avg)
- Black Arrestees:
   - Use of force ranges from 5.8% (2017) to 9.6% (2021) of arrestees
   - Average 2017–2019 = 5.9% arrestees with force
   - Average 2020–2022 = 8.1% arrestees with force
   - Increase of 37.3% arrestees with force (2017–2019 avg vs. 2020–2022 avg)
- Hispanic Arrestees
   - Use of force ranges from 4.5% (2019) to 7.1% (2021) of arrestees
   - Average 2017–2019 = 4.8% arrestees with force
   - Average 2020–2022 = 6.8% arrestees with force
   - Increase of 41.7% arrestees with force (2017–2019 avg vs. 2020–2022 avg)

In short, although Black arrestees were more likely to have force used against them, White arrestees experienced the largest increase in the likelihood of force.

TIME SERIES ANALYSES

Finally, it is instructive to consider the change in the percentage of arrestees who had force used against them over time as related to seminal events. **Figure 5.4** shows the monthly percentage of APD arrestees that had force used against them from 2017 to 2022. Again, there is a consistent pattern of proportional stability in arrestees who experienced force between January 2017 and March 2020. Beginning in April 2020 (following the onset of the COVID-19 pandemic social changes), the proportion of arrestees who had force used against them *doubled* to roughly 10% for that month. After October 2020, the ratio of arrestees who experienced force remained consistently higher than in the pre-COVID period. On average, 5.0% of arrested individuals had force used against them up until March 2020, while the post-March 2020 period accounted for the highest percentage of arrestees who experienced force (an average of 7.7% from April 2020 to December 2022, with a high of 12.1% of all arrests in January 2021).

Figure 5.4. Percentage of Arrested Individuals Who Experienced Use of Force by Month, 2017–2022 (n = 44,954)



The results of the interrupted time series models below show the change in the force counts while controlling for the number of arrest incidents (equating to a ratio of force counts per arrest) while accounting for seasonal monthly trends and previously established patterns.[48] **Table 5.3** shows that net of seasonality, the ratio of force counts to arrest counts increased by 108% in the post-COVID period. When the NPI team disaggregated the change in force by race, there was a statistically significant difference across all racial/ethnic groups on this rate of change in arrestees experiencing force. Specifically, counts of force (per arrest count) for White individuals increased by 120% compared to 83.3% and 80.7% increases for Black and Hispanic individuals, respectively.

Table 5.3. Interrupted Time Series Analyses for APD Use of Force Within Arrests, 2017-2022

	Total UoF	White UoF	Black UoF	Hispanic UoF
	B (SE) [Exp(B)-1]	B (SE) [Exp(B)-1]	B(SE) [Exp(B)-1]	B (SE) [Exp(B)-1]
Intercept	-2.73* (0.057)	-3.73* (0.093)	-3.50* (0.085)	-4.51* (0.142)
Post-COVID	0.738* (0.033) [1.08]	0.791* (0.057) [1.20]	0.606* (0.056) [.833]	0.592* (0.085) [0.807]

+All regression models include February – December monthly dummy variables (they are excluded from tables for parsimony). Count of arrests serves as the exposure variable (Coefficients are a rate change of arrests). *$p < 0.05$

48 The models include seasonal monthly dummy variables and control for the monthly count of arrests as the exposure variable. It treats the post-COVID change in arrest as the denominator (i.e., the exposure variable) with the change in uses of force among arrestees serving as the numerator (i.e., force per arrest), equating to an analysis of the change in ratios of force-within-arrests over time.

MULTIVARIATE RESULTS: USE OF FORCE WITHIN ARRESTS

As described in Section 2, multivariate models measure each predictor variable's individual and independent impact on the outcome (i.e., force) while holding all other variables in the model constant. This analysis allows the NPI team to isolate the impact of the key variable of interest – arrestee race/ethnicity – on force given similar characteristics of the incident, arrest, and person included in the model.

Table 5.4 below provides the descriptive statistics for the full arrest data[49] where the outcome of interest is use of force (0 = no force, 1 = force used). In approximately 39% of all encounters that resulted in arrest, the suspect had an outstanding warrant at the time of the arrest. For the six-year period, roughly 6% of all arrested individuals had force used against them, and 30% were arrested post-March 2020 (i.e., COVID-19 pandemic). Approximately 40% of all arrestees were Black, compared to 31% White and 26% Hispanic.

[49] The total number of arrestees = 44,954; however, 37 cases were removed from the analyses due to missing data on one or more of the variables included in the analyses.

Table 5.4. Descriptive Statistics for Model Predicting Use of Force Within Arrests (n = 44,917)

Variables	Mean	Standard Deviation	Minimum	Maximum
Force	0.06	0.234	0	1
Legal Characteristics				
Arrestee had outstanding warrant	0.39	0.487	0	1
Violent offense charge	0.07	0.260	0	1
Incident Characteristics				
Post-March 2020	0.30	0.458	0	1
Quarter 1	0.26	0.440	0	1
Quarter 2	0.25	0.432	0	1
Quarter 3	0.26	0.438	0	1
Quarter 4	0.23	0.421	0	1
Weekend	0.46	0.499	0	1
Nighttime	0.48	0.499	0	1
Multiple arrestees	0.10	0.302	0	1
Arrestee Characteristics				
Male	0.74	0.441	0	1
Age	32.7	11.56	< 1	83
White	0.31	0.445	0	1
Black	0.40	0.490	0	1
Hispanic	0.26	0.437	0	1
Other/Unknown	0.03	0.174	0	1

Given the sizable shift in arrests due to the onset of the COVID-19 pandemic – verified through the time series analyses in Section 3 – the NPI team conducted two regression analyses using different periods:

Period 1: Jan 1, 2017–Mar 31, 2020
Period 2: Apr 1, 2020–Dec 31, 2022

In essence, the NPI team split the data into the pre- and post-COVID periods to assess whether there were notable changes in the predicted probabilities of force within arrests by race/ethnicity after the number of arrests was essentially cut in half (post-COVID). **Table 5.5** below shows the results of statistical analyses predicting the factors that influence whether arrestees had force used against them after controlling for legal, incident, and arrestee characteristics.

First, Black arrestees were 1.37 times more likely than White arrestees to have force used against them pre-COVID, even after accounting for other factors. However, the strength of this effect (odds ratio) is substantively small in magnitude and decreases after the onset of COVID-19. During Period 2, Black arrestees were 1.17 times more likely than White arrestees to have force used against them. Hispanic arrestees, by contrast, did not differ significantly from White arrestees during Period 1. After the onset of COVID-19, Hispanic arrestees were 1.24 times less likely to experience force than White arrestees. Additionally, across both periods, males were 1.35 times more likely than females to have force used against them. Finally, if an individual was arrested with multiple arrestees in the same incident, that individual was 1.8 to 1.9 times more likely to have force used in the arrest, depending on whether the arrest occurred before or after the onset of COVID-19.

Multivariate analysis can only statistically control those variables that are measured. Specification error occurs due to the inability to specify all factors that might influence the outcome. If these unmeasured variables vary across racial/ethnic groups, their inclusion in the statistical models would increase or lessen the predicted impact of individuals' race/ethnicity on the likelihood of force. The interpretation of multivariate results must keep this limitation in mind.

Table 5.5. Logistic Regression Predicting Use of Force Within APD Arrests, 2017–2022 (n = 44,917)

	Period 1 (n = 31,497)		Period 2 (n = 13,420)	
Independent Variables	B (SE)	Odds Ratio	B (SE)	Odds Ratio
Intercept	-2.647* (0.125)	---	-1.907* (0.179)	---
Legal Characteristics				
Arrestee had outstanding warrant	-1.792 (0.087)	5.99	-1.091 (0.089)	2.98
Violent offense charge	0.010 (0.097)	--	0.337 (0.086)	1.40
Incident Characteristics				
2018	-0.027 (0.063)	---	---	---
2019	0.066 (0.065)	---	---	---
2021	---	---	0.142 (0.094)	---
2022	---	---	0.100 (0.090)	---
Quarter 2	0.002 (0.073)	---	-0.134 (0.103)	---
Quarter 3	0.073 (0.071)	---	-0.052 (0.101)	---
Quarter 4	-0.017 (0.075)	---	0.074 (0.102)	---
Weekend	-0.125* (0.053)	1.13	-0.010 (0.066)	---
Nighttime	0.423* (0.054)	1.53	0.006* (0.067)	---
Multiple arrestees	0.663* (0.068)	1.94	0.604* (0.095)	1.83
Arrestee Characteristics				
Male	0.298* (0.064)	1.35	0.190* (0.078)	1.21
Age	-0.017* (0.002)	1.02	-0.019* (0.003)	1.02
Black	0.311* (0.065)	1.37	0.156 (0.080)	1.17
Hispanic	-0.040 (0.076)	---	-0.215 (0.091)	1.24
Other/Unknown	0.091 (0.148)	---	-0.307 (0.222)	---
Model Fit Statistics				
	Nagelkerke R-Square value = 0.098		Nagelkerke R-Square value = 0.061	

*p < 0.05; only statistically significant odds ratios are presented. Odds ratios for negative coefficients are calculated as 1/expB

Predicted Probability of Force from Regressions

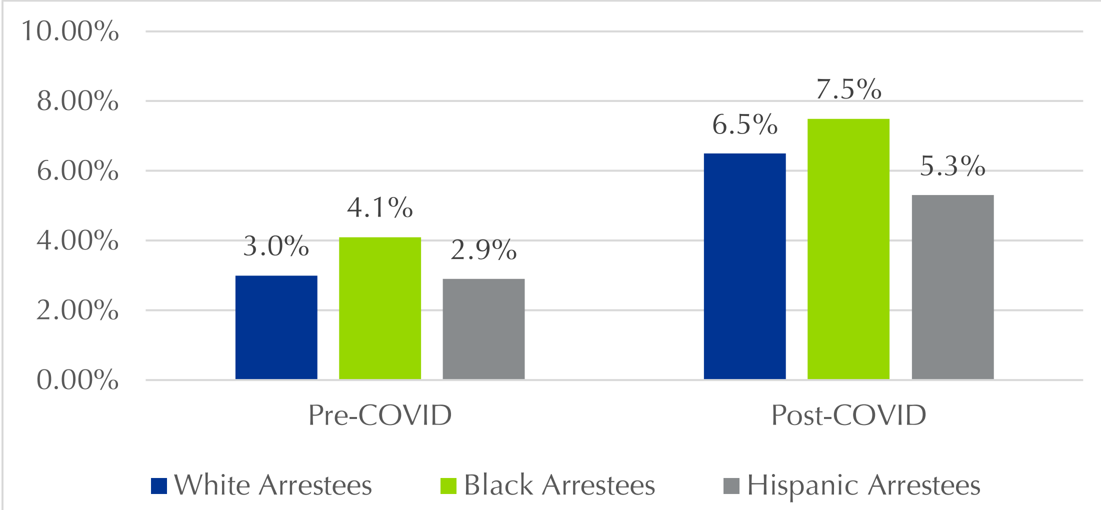
While the odds ratios (displayed in the table above) can describe the strength of a measure relative to other variables in the model, predicted probabilities are a more precise estimation method that demonstrates the impact of the independent variables in a regression model. A predicted probability is simply the probability of an event occurring; in this case, the probability that an individual is involved in police use of force in the arrest.[50]

Figure 5.5 shows the predicted probability of force being used in an arrest based on the arrestee's demographic characteristics, net of all other factors in the model. There are three noteworthy findings regarding the likelihood of force within arrests by race/ethnicity in the period when arrests were more commonplace (Model 1, pre-COVID) and when they were considerably restricted (Model 2, post-COVID).

- The probability of force being used during arrests increased over time.
- As arrest counts declined, partly due to reduced arrests for less serious offenses, the probability that force was used increased for all racial/ethnic groups.
    - White arrestees 3.0% pre-COVID, 6.5% post-COVID
    - Black arrestees 4.1% pre-COVID, 7.5% post-COVID
    - Hispanic arrestees 2.9% pre-COVID, 5.3% post-COVID
- The differences in the probability of force being used during arrests across racial/ethnic groups (i.e., the differences across the groups relative to each other) were cut in half. The remaining differences in probability of experiencing force during arrest were reduced considerably post-COVID.

[50] The baseline predicted probability is the foundation of the regression model, where all estimates are set to their average values. To determine the effect size of statistically significant independent variables, the average values are changed to the low-to-high values of the measures – which can be interpreted as, "all else being equal in the model, the likelihood that x is associated with y" is demonstrated by a given predicted probability.

Figure 5.5. Predicted Probability of Force During Arrests in Models 1 & 2 (Pre- & Post-COVID)



SECTION SUMMARY

The analyses in this section examined use of force within arrests, as use of force is most common during situations that involve arrests. These analyses are based on a sample of 44,954 arrested individuals to better understand why 2,608 of those arrestees experienced force. The NPI team compared the percentage of arrestees that had force used against them over time, by racial/ethnic groups, and used multivariate statistical analyses to better understand what factors predict whether force is used against arrested individuals.

On average, approximately 5.8% of arrested individuals also had force used against them. Although *fewer* individuals had force used against them from 2020 to 2022 compared to 2017 to 2019, a *larger* percentage of arrestees had force used against them because of the post-COVID decline in the overall number of arrestees. Accounting for seasonality, the ratio of force counts to arrest counts increased by 108% in the post-COVID period. This increase was unviersal across all racial/ethnic groups but demonstrated a larger, statistically significant increase for White arrestees (120%) than for their Black (83.3%) and Hispanic (80.7%) counterparts.

Multivariate analyses show small to marginal disparities in use of force for Black arrestees compared to White arrestees. Still, these disparities are smaller after March 2020 than in the period before COVID-19 (odds ratios = 1.37 pre-COVID, 1.17 post-COVID). Hispanic arrestees were not more likely to experience use of force than White arrestees in either period and were 1.2 times significantly less likely to have force used against them post-COVID.

SECTION 6: CONCLUSION

To support its monitorship of the Colorado Attorney General's Office consent decree with the City of Aurora, IntegrAssure engaged the National Policing Institute (NPI) to analyze and interpret enforcement data from the Aurora Police Department to develop baseline measures that may be used to examine racial disparities in police activity and outcomes over time. This report presents the findings from the NPI team's examination of the patterns and trends in the APD's criminal summonses, arrests, and use of force reported from 2017 to 2022 to inform future analyses. This section summarizes the main findings of the report and provides recommendations for IntegrAssure and the APD to support comprehensive data collection of the APD's enforcement activities and implement policies and training to promote community and officer safety.

KEY FINDINGS

The NPI team conducted a series of statistical analyses to understand APD's enforcement activities better. The key findings are summarized below.

- Crime, especially serious and violent crime, steadily increased from 2017 to 2022 in the City of Aurora. There has been a 20% increase for all criminal offenses from 2017 to 2022. When serious crime is considered, Uniform Crime Reports (UCR) Part I crimes (murder, rape, robbery, aggravated assault, burglary, larceny, auto theft, and arson) have increased 44% over the past six years, and violent crime has risen 82%. Time series analyses indicate a consistent upward trend of reported crime that was not significantly reduced or accelerated by seminal events, including the onset of the COVID-19 pandemic.

- As crime continued to increase from 2017–2022, the number of criminal summonses and arrests significantly decreased. This decline in enforcement activity was accelerated by the onset of the COVID-19 pandemic in March 2020, resulting in significant reductions in the use of criminal summonses (11.2% decline) and arrests (approximately 50% decline) that continued through 2022.

- The number of subjects who had force used against them by police was relatively stable across the six-year period. This pattern was not interrupted by the COVID-19 pandemic or any other seminal event examined. However, as arrests declined, the percentage of arrestees who experienced police force significantly increased post-COVID. On average, 5.0% of arrested individuals experienced force until March 2020, while the average from April 2020 to December 2022 was 7.7%.

- Limitations associated with the APD's use of force data restricted the NPI team's ability to conduct more in-depth analyses of patterns and trends that might explain the stability in use of force despite the decline in arrests and summonses.

- Several different analytic approaches were used to estimate the presence and level of racial/ethnic disparities in APD arrests and uses of force, including both benchmark and multivariate regression models. Combined, these findings suggest small to marginal disparities in arrests and uses of force for Black subjects when compared to White subjects. For Hispanic subjects, small to marginal disparities in arrests were evident, but there were no disparities in use of force for Hispanic subjects when compared to their White counterparts.

- Of the 44,954 individuals arrested by APD officers, 40.3% were Black, 30.8% were White, 25.8% were Hispanic, and 3.1% were of other or unknown racial/ethnic backgrounds. *Arrest benchmark analyses* compared these percentages to four comparison populations: (1) residential population, (2) all crime suspects, (3) crime suspects for Part I offenses, and (4) crime suspects for Part I violent offenses.
    - These analyses show small to marginal (DR=1.06 to 1.33) or no (DR=0.70-0.84) racial/ethnic disparities for Black and Hispanic individuals in arrests using non-census benchmark comparisons.
    - For the most recent period (post-COVID), racial/ethnic disparities decreased, and some suspect-based benchmarks showed that Black and Hispanic individuals were less likely to be arrested than White individuals.
    - The residential population benchmarks produced a disparity ratio of 3.43 for Black individuals and 1.26 for Hispanic individuals. The validity of this benchmark (as an accurate measure of the population at risk of arrest) has been widely questioned and debunked by many experts.

- Of the 3,783 individuals who had force used against them, 43.1% were Black, 33.5% were White, 15.3% were Hispanic, 5.7% were of unknown race/ethnicity, and 2.5% were other racial/ethnic backgrounds. *Use of force benchmark analyses* compared these percentages to eight comparison populations: (1) residential population, (2) criminal summonses, (3) all arrestees, (4) arrestees for

Part I offenses, (5) arrestees for Part I violent offenses, (6) all crime suspects, (7) crime suspects for Part I offenses, and (8) crime suspects for Part I violent offenses.

- The NPI team found substantively small (DR=1.04-1.22) and, in some cases, no disparities (DRs less than 1.0) in use of force for Black individuals when using non-census benchmark comparisons.
- After the onset of COVID-19, these small to marginal disparities were further reduced or eliminated across all benchmarks.
- As with arrests, only the residential population benchmark demonstrated racial/ethnic disparities in police use of force, and only for Black compared to White individuals.
- No disparities in use of force for Hispanic individuals were evident across all benchmarks either before or after the onset of COVID-19 (all DRs less than 1.0).

- *Multivariate analyses* were used to explore the factors that influence whether arrestees experience force. The results of the multivariate analyses must be interpreted cautiously because the strongest known predictors of use of force (e.g., suspect resistance, intoxication, presence of a weapon, etc.) could not be included in the statistical models.
  - These analyses show that Black arrestees were significantly more likely to have force used against them compared to White arrestees after controlling for other situational, legal, and arrestee characteristics. Although the differences in the likelihood of use of force for Black compared to White arrestees is statistically significant, it represents a substantively small difference in the predicted probabilities of use of force (4.1% for Black arrestees vs. 3.0% for White arrestees pre-COVID and 7.5% for Black arrestees and 6.5% for White arrestees post-COVID).
  - Furthermore, the racial differences are smaller after March 2020 than pre-COVID (odds ratios=1.37 pre-, 1.17 post).
  - The multivariate analyses also show that Hispanic arrestees were *not* significantly more likely to experience force than White arrestees during the six-year period after controlling for other situational, legal, and arrestee characteristics.
  - Post-COVID, Hispanic arrestees were 1.2 times significantly *less* likely to experience force post-COVID than White arrestees.
  - The differences in the probability of force within arrests across racial/ethnic groups were cut in half as White arrestees' probability of

force increased more than Black and Hispanic arrestees' probability of force post-COVID.

RECOMMENDATIONS

Based on the findings reported above, the NPI team recommends five primary actions to support improvements to APD policy, training, and supervision.

Recommendation 1: Continue data collection system overhaul.

Before NPI's work with the APD, the department had already recognized the limitations of its use of force reporting system. APD has been actively developing a new system for reporting and collecting use of force data that should be operational soon. Unfortunately, the NPI team was reliant on historical use of force data to establish patterns and trends, and the available data limited the NPI team's analyses. The APD has been actively developing a new system for reporting and collecting use of force data that should be operational soon. Improvements to the reporting system will assist in better understanding the dynamics of use of force interactions, exploring whether there are racial/ethnic differences in correlates of use of force, and examining the factors that predict subject and officer injuries, all of which can potentially inform additional improvements to use of force policy and training.

The limitations to the use of force data included problems with the reliability and validity of existing data fields and the failure to capture key information on APD officers' use of force in both arrest and use of force reporting systems. The APD's use of force data would be greatly improved by expanding the data fields collected within the use of force report (e.g., subject resistance) and improving the reliability and validity of the data captured within the existing fields (see Recommendation 2). The NPI team has reviewed and provided recommendations to the APD's working draft of an updated use of force report. However, the APD should also review the Police Executive Research Forum's (PERF) Use-of-Force Data Framework for a comprehensive list of data fields to consider including.[51]

APD personnel responsible for enhancing the use of force reporting system should carefully review the limitations in the data collected that are noted throughout this report, paying particular attention to the system's ability to ensure data fields are collected at the appropriate unit of analysis. For example, reason for force, type of force, and injuries would ideally be connected to each officer's use of force against each subject. This link between officers and subjects is critical for in-depth analyses of types of force (and their effectiveness) and officer and subject injuries.

[51] See: https://www.policeforum.org/assets/PERFUOFDataFramework.xlsx

The APD should also consider adding data fields to the arrest reporting system to understand the factors influencing whether officers use force during arrests. Although arrest reports are completed based on administrative and legal requirements, the addition of a small number of key data fields (e.g., subject resistance, whether a weapon was present, and whether an individual was impaired) would assist greatly in the understanding of officer decision-making related to use of force.

Recommendation 2: Add accountability checks for accurate data collection to demonstrate its importance.

For APD to continue to be data-driven in its practices and to provide transparency to the community, the department must improve the quality of its use of force data. As APD is developing its new use of force data collection system, care should be taken to develop or enhance reliability and validity checks, including validation measures within the data reporting system, APD's chain of command review processes, and periodic data audits.

In the NPI team's experience (Engel et al., 2023), law enforcement agencies can make dramatic improvements in missing data and logical inconsistencies by setting up the reporting system to:

- use drop-down categorical menus where appropriate,
- open certain data fields only when needed,
- make certain data fields mandatory,
- warn personnel of possible data entry errors in the report before submission.

These validation checks are illustrated using the injury data fields as an example:

- To minimize missing data on whether a subject was injured, the reporting system should be set to mandate a valid response (yes, no, unknown) for injury or warn officers when the field lacks a valid response.
- The injury nature field should be set only to open when the response to "subject injured" is yes.
- If there is an injury, having categories of injuries (e.g., abrasions, TASER probes, fracture, etc.) to select from would provide some uniformity to the injury nature field[52] that would make coding the injury nature variable far less cumbersome and facilitate injury type and severity analyses.

Following the completion of use of force reports, reviewers in the chain of command should ensure that all necessary data fields are completed and send them back for

[52] For the six years of provided data, there were 1,100 different responses for injury nature.

corrections as needed. APD should consider periodic data audits of the various data collection systems, especially for use of force, to check for inaccuracies and maintain quality control. Including these measures to improve the quality of data collection will reinforce to personnel completing use of force reports that accuracy and completeness in reporting are essential.

Recommendation 3: Continue updates in use of force policy and training.

The APD is already in the process of revising (and renumbering) use of force-related policies.[53] As part of these updates and as previously recommended in the *Technical Report*, APD should consider revising Directive 05.05 *Reporting Use of Force*[54] to reclassify the pointing of a firearm from Tier Zero to Tier One.[55] This would facilitate more detailed reporting and evaluation by supervisors and commanders to ensure these actions are in line with department policy and reduce the risk of accidental or unjustified shootings.[56]

Under existing APD policies, all levels of force have associated reporting requirements, each with detailed instructions on recording the event and the required phases of supervisory review. Based on recent policy updates, APD Directives 05.05 and 05.06 direct that any uses of Tier Zero, One, or Two types of force require the officer who used that force to complete a Contact Data Collection (CDC) Report in the Benchmark System. For Tier Zero, if there is no associated CAD call, the officer must create a CAD call, notify their supervisor, and complete the CDC form. Importantly, this results in differences in the information collected for Tier Zero since Tiers One and Two have additional reporting requirements. It is unknown to the NPI team how the APD plans to analyze the use of force information collected via the CDC report compared to the use of force reports. Regardless of whether pointing of a firearm remains a Tier Zero or becomes a Tier One reportable force, it is recommended that the available data on the use of pointing of firearms be analyzed and reviewed regularly.

In 2023, the APD trained its personnel using the PERF's de-escalation training: *Integrating, Communications, Assessment, and Tactics (ICAT)*. Research evaluating the

[53] https://www.auroramonitor.org/_files/ugd/074938_7218e294cc8547e19dd325af72875a55.pdf
https://public.powerdms.com/AURORAPD/tree

[54] This was formerly 05.04 Reporting and Investigating the Use of Tools, Weapons, and Physical Force.

[55] Since January 1, 2016, the APD has classified the pointing of a firearm as a Tier Zero type of force; this level of force is described by department policy as a "display of force." DM 05.05 *Reporting and Investigating the Use of Tools, Weapons, and Physical Force*
https://public.powerdms.com/AURORAPD/tree/documents/107

[56] Notably, recent research suggests that police agencies with policies requiring documentation of pointing of a firearm have significantly lower rates of officer-involved shootings. This policy was not associated with increased injury or death rates among officers (Jennings & Rubado, 2017; Shjarback et al., 2021).

ICAT training demonstrated significant reductions in officer use of force and community member and officer injuries (Engel et al., 2022). PERF recently published an ICAT training implementation guide for agencies with several strategies for maximizing and sustaining the benefits of de-escalation training (PERF, 2023). The NPI team recommends that the APD continue to implement these and other evidence-based approaches.

Recommendation 4: Continue to track changes in racial/ethnic disparities in APD enforcement actions using multiple measures and analytical techniques.

Determining whether racial/ethnic disparities exist in enforcement actions can be complex. Nevertheless, understanding the extent to which disparities exist and under what circumstances can provide critical information to guide any law enforcement agency's approach to addressing them. The current report provides valuable baseline measures for trends in crime, enforcement outcomes, and racial/ethnic disparities from 2017 to 2022.

IntegrAssure, in their role as the Independent Consent Decree Monitor, can use the information provided in this report to aid in their ongoing assessments of whether the City has changed "in measurable ways, how Aurora Police engages with all members of the community, including by reducing any racial disparities..." (Consent Decree, 2022, p.7). The APD should also use this information to establish their own measures and expectations for performance and enforcement operations to ensure the department meets consent decree mandates and adopts best practices.

The APD should continue to monitor trends in enforcement and racial/ethnic disparities with additional years of data as it becomes available. The APD should also begin regular analysis of the Contact Data Collection forms, which were initiated in July 2022 to document all enforcement or investigatory interactions with the public. A comprehensive understanding of enforcement patterns and trends requires analysis of multiple data sources and statistical techniques. In addition, a totality of the circumstances approach to understanding racial/ethnic disparities in enforcement should incorporate the perspectives of multiple stakeholders within the department, City of Aurora leadership, and community members. Therefore, the APD should consider partnering with an independent research team to continue this work.

These quantitative and qualitative data can increase understanding of the factors influencing police enforcement actions, the role of race/ethnicity, and strategies to ensure fair and impartial policing in all encounters with the public.

Recommendation 5: Implement effective and equitable crime reduction strategies immediately – especially focused on violence – and continually monitor the impact on reported crime, enforcement disparities, and community sentiment.

The findings related to crime trends in Aurora indicate that there was a substantial increase in Part I offenses and violent crime from 2017 to 2022. The time series analyses highlighted that, although COVID-19 did not impact this crime trend, it significantly reduced the overall number of criminal summonses and arrests by the APD. At the same time, disparities in arrests and uses of force decreased for Black and Hispanic individuals in the post-COVID period. Thus, it is critical for the APD and the City of Aurora to implement strategies that can effectively balance violent crime reductions while maintaining the progress that has been achieved in reducing racially disparate outcomes.

In the last 10-15 years, several evidence-based strategies have proven to be effective at reducing violent crime while avoiding exacerbating racial disparities (McManus et al., 2020). In particular, it is important to recognize that violence is highly concentrated among a small number of people and places, often as a result of historical underinvestment and neglect. Many promising violence reduction strategies focus on those two elements specifically, and for most cities, a combination approach is the most effective.

Some effective place-based strategies include Place-Network Investigation (Herold et al., 2020), hot-spots policing (Braga et al., 2019; Corsaro et al., 2021; Weisburd et al., 2022), cleaning/greening vacant lots (Branas et al., 2018; Sadatsafavi et al., 2022), abandoned buildings remediation (Kondo et al., 2015; Jay et al., 2019; South et al., 2021), improved street lighting (Chalfin, 2021; Mitre-Becerril et al., 2022), and community reinvestment (Culyba et al., 2016; Kondo et al., 2018; Sharkey, 2018).

In addition, it is important to focus on those individuals at the highest risk of violent victimization or commission by using strategies such as street outreach and violence interruption programs (Buggs et al., 2021; Roman et al., 2017; Webster et al., 2013), hospital-based violence intervention programs (Affinati et al., 2016; Bell et al., 2018; Purtle et al., 2013), employment programming (Heller et al., 2017; Bhatt et al., 2023), and focused deterrence strategies (Braga et al., 2018; Corsaro & Engel, 2015; Engel et al., 2013).

Moving towards a comprehensive, city-wide violence prevention strategy that uses evidence-based strategies focusing on the highest-risk people and places would help Aurora reduce violence while maintaining the positive improvement in racial disparities in policing outcomes.

CONCLUSION

In conclusion, the findings presented within this report identify critical baseline measures that may be used to compare patterns, trends, and outcomes associated with the Aurora Police Department's enforcement activities over time. In examining racial/ethnic disparities, the present analyses suggest that differences in the APD's enforcement actions across racial/ethnic groups are statistically small and decreasing over time. However, the methodological and data quality limitations affecting these analyses warrant caution in interpreting these findings. It is important to note that, regardless of the available data or statistical analyses employed, the aggregate, quantitative examination of patterns and trends in enforcement outcomes *cannot* determine whether racial bias is the source of the differences observed in APD officers' enforcement actions. As such, the information presented within this report is best used to establish measures that may be examined over time to identify patterns and trends in APD enforcement activities and assess changes in policing outcomes as additional reforms are implemented to align with consent decree mandates.

Pairing continuous assessment with the implementation of reforms can support the APD in building evidence around the impact of their practices and inform alterations to training, policy, and protocols (as appropriate) to achieve desired outcomes.

REFERENCES

Affinati, S., Patton, D., Hansen, L., Ranney, M., Christmas, A., Violano, P., Sodhi, A., Robinson, B., Crandall, M. (2016). Hospital-based violence intervention programs targeting adult populations: An eastern association for the surgery of trauma evidence-based review. *Trauma Surgery & Acute Care Open*, doi:10.1136/tsaco-2016-000024

Alpert, G. P., & Dunham, R. (1999). The force factor: Measuring and assessing police use of force and suspect resistance. Report to the National Institute of Justice. Use of force by police: Overview of national and local data. Washington, DC: U.S. Department of Justice.

Alpert, G. P., Smith, M. R., & Dunham, R.G. (2004). Toward a better benchmark: Assessing the utility of not-at-fault traffic crash data in racial profiling research. *Justice Research and Policy*, 6, 43–69.

Bell, T. M., Gilyan, D., Moore, B. A., Martin, J., Ogbemudia, B., McLaughlin, B. E., ... & Zarzaur, B. L. (2018). Long-term evaluation of a hospital-based violence intervention program using a regional health information exchange. *The journal of trauma and acute care surgery,* 84(1), 175.

Betensky, R. (2019). The p-Value Requires Context, Not a Threshold. *The American Statistician*, 73(1), 115-117. DOI: 10.1080/00031305.2018.1529624

Bhatt, M. P., Heller, S. B., Kapustin, M., Bertrand, M., & Blattman, C. (2023). *Predicting and preventing gun violence: An experimental evaluation of READI Chicago* (No. w30852). National Bureau of Economic Research.

Braga, A. A., Turchan, B. S., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing and crime reduction: An update of an ongoing systematic review and meta-analysis. *Journal of experimental criminology*, 15, 289-311.

Braga, A. A., Weisburd, D., & Turchan, B. (2018). Focused deterrence strategies and crime control: An updated systematic review and meta-analysis of the empirical evidence. *Criminology & Public Policy*, 17(1), 205-250.

Branas, C. C., South, E., Kondo, M. C., Hohl, B. C., Bourgois, P., Wiebe, D. J., & MacDonald, J. M. (2018). Citywide cluster randomized trial to restore blighted vacant land and its effects on violence, crime, and fear. *Proceedings of the National Academy of Sciences*, 115(12), 2946-2951.

Brown, R., Engel, R.S., Cherkauskas, J., Corsaro, N., & Kurtz, J. D. (2022). Assessment of Colorado Springs Police Department Use of Force. Report submitted to the Colorado Springs Police Department, Office of the Chief. https://coloradosprings.gov/sites/default/files/inline-images/cspd_use_of_force_final_transparency_matters_report_april_2022.pdf

Buggs, S. A., Webster, D. W., & Crifasi, C. K. (2022). Using synthetic control methodology to estimate effects of a Cure Violence intervention in Baltimore, Maryland. *Injury prevention*, 28(1), 61-67.

Cesario, J., Johnson, D. J., & Terrill, W. (2019). Is There Evidence of Racial Disparity in Police Use of Deadly Force? Analyses of Officer-Involved Fatal Shootings in 2015–2016. *Social psychological and personality science*, *10*(5), 194855061877510.

Chalfin, A., Hansen, B., Lerner, J., & Parker, L. (2021). Reducing crime through environmental design: Evidence from a randomized experiment of street lighting in New York City. *Journal of Quantitative Criminology*, 1-31.

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics—simulation and Computation, 39(4), 860-864*.

Consent Decree & Judgment, (November 16, 2021). State of Colorado, ex rel. Philip J. Weiser, Attorney General v. City of Aurora, Colorado (Arapahoe County District Court, Colorado).

Corsaro, N. & Engel, R. S. (2015). The most challenging of contexts: Assessing the impact of focused deterrence on serious violence in New Orleans. *Criminology and Public Policy*, 14(3): 471-505.

Corsaro, N., Engel, R. S., Herold, T. D., Yildirim, M., & Motz, R. T. (2021). Hot spot policing in Las Vegas: Results from a blocked randomized controlled trial in chronic violent crime locations. *Journal of Experimental Criminology*, 1-23. https://doi.org/10.1007/s11292-021-09485-8.

Culyba, A. J., Jacoby, S. F., Richmond, T. S., Fein, J. A., Hohl, B. C., & Branas, C. C. (2016). Modifiable neighborhood features associated with adolescent homicide. *JAMA Pediatrics*, 170(5), 473-480.

Davis, E., Whyde, A., & Langton, L. (2018). *Contacts between police and the public, 2015*. Washington, DC: U. S. Department of Justice.

Engel, R. S. (2015). Police encounters with people with mental illness: Use of force, injuries, and perceptions of dangerousness, *Criminology and Public Policy*, 14(2): 247-251.

Engel, R., & Calnon, J. (2004). Comparing Benchmark Methodologies for Police-Citizen Contacts: Traffic Stop Data Collection for the Pennsylvania State Police. *Police Quarterly*, *7*(1), 97–125.

Engel, R.S., Cherkauskas, J.C., Corsaro, N., & Yildirim, M. (2023). *2022 Pennsylvania State Police Traffic Stop Study: January 1 – December 31, 2022*. Report submitted to the Commissioner of the Pennsylvania State Police.

Engel, R. S., Corsaro, N., Isaza, G. T., & McManus, H. D. (2022). Assessing the impact of de-escalation training on police behavior: Reducing police use of force in the Louisville, KY Metro Police Department. *Criminology & Public Policy*.

Engel, R. S., Corsaro, N., Isaza, G. T., & McManus, H. D. (2020). *Examining the impact of Integrating Communications, Assessment, and Tactics (ICAT) de-escalation training for the Louisville Metro Police Department: Initial findings.* Alexandria, VA: International Association of Chiefs of Police. Retrieved from https://www.theiacp.org/sites/default/files/Research%20Center/LMPD_ICAT%20Evaluation%20Initial%20Findings%20Report_FINAL%2009212020.pdf

Engel, R. S., Tillyer, M. S., & Corsaro, N. (2013). Reducing gang violence using focused deterrence: Evaluating the Cincinnati Initiative to Reduce Violence (CIRV). *Justice Quarterly*, 30(3): 403-439.

Engel, R. S., Sobol, J. J., & Worden, R. E. (2000). Further exploration of the demeanor hypothesis: The interaction effects of suspects' characteristics and demeanor on police behavior. Justice Quarterly, 17(2), 235–258.

Engel, R. S., & Swartz, K. (2014). Race, crime, and policing. In The Oxford handbook of ethnicity, crime, and immigration, eds. Sandra M. Bucerius and Michael Tonry, 135–65. New York, NY: Oxford University Press.

Fenton, N., Neil, M., & Frazier, S. (2020). The role of collider bias in understanding statistics on racially biased policing. *arXiv preprint arXiv:2007.08406*.

Fridell, L. A. (2004). *By the Numbers: A Guide for Analyzing Race Data from Vehicle Stops*. Washington, DC: Police Executive Research Forum.

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

Garner, J. H., Maxwell, C. D., & Heraux, C. G. (2002). Characteristics associated with the prevalence and severity of force used by the police. *Justice Quarterly*, *19*(4), 705–746.

Garner, J. H., Hickman, M. J., Malega, R. W., Maxwell, C. D. (2018). Progress toward national estimates of police use of force. *PLoS ONE*, *13*(2), e0192932.

Geller, A., Goff, P. A., Lloyd, T., , A., Obermark, D., & Glaser, J. (2021). Measuring racial disparities in police use of force: methods matter. *Journal of Quantitative Criminology*.

Hanushek, E. A., & Jackson, J. E. (1977). *Statistical methods for social scientists*. Orlando, FL: Academic Press.

Heller, S., Pollack, H. A., & Davis, J. M. (2017). *The effects of summer jobs on youth violence*. National Criminal Justice Reference Service, Office of Justice Programs.

Herold, T. D., Engel, R. S., Corsaro, N., & Clouse, S. L. (2020). Place network investigations in Las Vegas, Nevada: Program review and process evaluation. https://www.theiacp.org/sites/default/files/Research%20Center/LVMPD_PIVOT%20Program%20Review_Final.pdf

Hickman, M. J., Piquero, A. R. & Garner, J. H. (2008). Toward a national estimate of police use of nonlethal force. *Criminology & Public Policy, 7*(4), 563-604.

Hudson, J., Fielding, S., & Ramsay, C. R. (2019). Methodology & reporting characteristics of studies using interrupted time series design in healthcare. *BMC Medical Research Methodology, 19*(1), 137.

Jay, J., Miratrix, L.W., Branas, C.C., Zimmerman, M.A., & Hemenway, D. (2019). Urban building demolitions, firearm violence and drug crime. *Journal of Behavioral Medicine*, 42(4), 626-634.

Jennings, J. T., & Rubado, M. E. (2017). Preventing the use of deadly force: The relationship between police agency policies and rates of officer-involved gun deaths. *Public Administration Review*, *77*(2), 217-226.

Jung, J., Corbett-Davies, S., Shroff, R., & Goel, S. (2018). Omitted and included variable bias in tests for disparate impact. *arXiv preprint arXiv:1809.05651*.

Klinger, D. A., & Bridges, G. S. (1997). Measurement error in calls-for-service as an indicator of crime. *Criminology*, *35*(4), 705-726).

Knox, D., W. Lowe, and J. Mummolo. (2020a). Administrative Records Mask

Racially Biased Policing. *American Political Science Review*. https://www.cambridge.org/core/journals/american-political-science-review/article/administrative-records-mask-racially-biased-policing/66BC0F9998543868BB20F241796B79B8

Knox, D., Lowe, W., & Mummolo, J. (2020b). Can Racial Bias in Policing Be Credibly Estimated Using Data Contaminated by Post-Treatment Selection?. *Available at SSRN 3940802*.

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*(3), 341-378.

Kondo, M. C., Andreyeva, E., South, E. C., MacDonald, J. M., & Branas, C. C. (2018). Neighborhood interventions to reduce violence. *Annual review of public health*, 39, 253-271

Kondo, M.C., Danya K., Bernadette C.H., MacDonald, J.M., & Branas, C.C. (2015). A difference-in-differences study of the effects of a new abandoned building remediation strategy on safety. *PloS One*, 10(7)

Kramer, R. & Remster, B. (2018). Stop, frisk, and assault? Racial disparities in police use of force investigatory stops. *Law & Society Review*, 52, 960-993.

Long J. S. (1997). Regression models for categorical and limited dependent variables. *Sage Advanced Quantitative Techniques in the Social Sciences Series*. Vol. 7. Thousand Oaks, CA: Sage.

Marvell, T. B., & Moody, C. E. (1996). Specification problems, police levels, and crime rates. *Criminology*, *34*(4), 609-646.

McManus, H., Engel, R. S., Cherkauskas, J.C., Light, S.C., Shoulberg, A.M. (2020). Street Violence Crime Reduction Strategies: A Review of the Literature. Report submitted to International Association of Chiefs of Police and Arnold Ventures.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2016). Applied multivariate research: Design and interpretation. Sage publications.

Mitre-Becerril, D., Tahamont, S., Lerner, J., & Chalfin, A. (2022). Can deterrence persist? Long-term evidence from a randomized experiment in street lighting. *Criminology & Public Policy*, 21(4), 865-891.

Morabito, M. S., Socia, K., Wik, A., & Fisher, W. (2017). The nature and extent of police use of force in encounters with people with behavioral health disorders. *International Journal of Law and Psychiatry*, 50, 31-37.

Police Executive Research Forum. (2023). *Implementing the ICAT Training Program at Your Agency*. https://www.policeforum.org/assets/ICATImplementation.pdf

Police Executive Research Forum. (2021). *What Police Chiefs and Sheriffs Need to Know About Collecting and Analyzing Use-of-Force Data*. https://www.policeforum.org/assets/CollectingAnalyzingUOFData.pdf

Police Executive Research Forum. (2016). Guiding principles on use of force. *Critical Issues in Policing Series*, 1-136.

Purtle, J., Dicker, R., Cooper, C., Corbin, T., Greene, M. B., Marks, A., Creaser, D., Topp, D., & Moreland, D. (2013). Hospital-based violence intervention programs save lives and money. *Journal of Trauma and Acute Care Surgery*, 75(2), 331-333.

Ratcliffe, J. H. (2004). Geocoding crime and a first estimate of a minimum acceptable hit rate. *International Journal of Geographical Information Science*, *18*(1), 61-72.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Sage.

Roman, C.G., Klein, H., Wolff, K.T., Bellamy, M.D., Reeves, K. (2017). *Philadelphia CeaseFire: Findings from the Impact Evaluation*. Temple University. https://cvg.org/wp-content/uploads/2020/03/SummaryofPhilaCeaseFireFindingsFormatted_Jan2017.pdf

Sadatsafavi, H., Sachs, N.A., Shepley, M.M., Kondo, M.C., Barankevich, R.A. (2022). Vacant lot remediation and firearm violence – A meta-analysis and benefit-to-cost evaluation. *Landscape and Urban Planning*. https://doi.org/10.1016/j.landurbplan.2021.104281

Sharkey, P. (2018). *Uneasy peace: The great crime decline, the renewal of city life, and the next war on violence*. WW Norton & Company.

Shjarback, J. A., White, M. D., & Bishopp, S. A. (2021). Can police shootings be reduced by requiring officers to document when they point firearms at citizens? *Injury prevention*, *27*(6), 508-513.

South, E. C., MacDonald, J., & Reina, V. (2021). Association between structural housing repairs for low-income homeowners and neighborhood crime. *JAMA network open*, 4(7), e2117067-e2117067.

Smith, M. R., Tillyer, R., & Engel, R. S. (2022). Race and the use of force by police revisited: Post-Ferguson findings from a large county police agency. *Police Quarterly*, 10986111221139442.

Smith, J. M., Tillyer, R., Engel, R. S., & Cherkauskas, J. C. (2019). *A Multi-Method Investigation of Officer Decision-Making and Force Used or Avoided in Arrest Situations: Tulsa, Oklahoma Police Department Administrative Data Analysis Report*. Cincinnati, OH: IACP/UC Center for Police Research and Policy.

Stroshine, M. S, & Brandl, S. G. (2019). The use, effectiveness, and hazards associated with police use of force: The unique case of weaponless physical force. *Police Practice and Research*, 1-18.

Terrill, W., & Mastrofski, S. D. (2002). Situational and officer-based determinants of police coercion. *Justice Quarterly*, 19(2), 215-248.

Tillyer, R., Engel, R. S., & Calnon Cherkauskas, J. (2010). Best practices in vehicle stop data collection and analysis. *Policing*, *33*(1), 69.

Webster, D. W., Whitehill, J. M., Vernick, J. S., & Curriero, F. C. (2013). Effects of Baltimore's Safe Streets Program on gun violence: A replication of Chicago's CeaseFire Program. *Journal of Urban Health*, 90, 27-40.

Weiser, P. (2021). *Investigation of the Aurora Police Department and Aurora Fire Rescue*. Denver: Colorado Office of the Attorney General.

Witte, R. S., & Witte, J. S. (2015). *Statistics*. John Wiley & Sons.

APPENDIX: SUPPLEMENTARY TABLES

Table A.1. Comparison of APD Use of Force Racial/Ethnic Disparity Ratios Across Benchmarks Pre & Post COVID

	Pre/Post COVID		Percent Race/Ethnicity			Disproportionality Indices			Disparity Ratios	
			White	Black	Hispanic	White	Black	Hispanic	Black	Hispanic
% Use of Force	Pre	(N=1902)	32.8% (658)	45.7% (916)	16.4% (328)	--	--	--	--	--
	Post	(N=1572)	34.2% (609)	40.1% (713)	14.1% (250)	--	--	--	--	--
Benchmark 1: % Residential Population[57]	Pre		43.5%	16.6%	29.0%	0.75	2.75	0.57	**3.65**	**0.75**
	Post		43.5%	16.6%	29.0%	0.79	2.42	0.49	**3.07**	**0.62**
Benchmark 2: Criminal Summons Population	Pre		32.6%	35.8%	25.8%	1.01	1.28	0.64	**1.27**	**0.63**
	Post		35.1%	37.3%	25.0%	0.97	1.08	0.56	**1.10**	**0.58**
Benchmark 3: % Arrestee Population (All crimes)	Pre		31.3%	40.7%	24.9%	1.05	1.12	0.66	**1.07**	**0.63**
	Post		29.6%	39.6%	27.9%	1.16	1.01	0.51	**0.88**	**0.44**
Benchmark 4: % Arrestee Population (Part I Crimes)	Pre		30.1%	43.4%	23.3%	1.09	1.05	0.70	**0.97**	**0.65**
	Post		28.7%	42.8%	26.0%	1.19	0.94	0.54	**0.79**	**0.46**
Benchmark 5: % Arrestee Population (Part I Violent Crimes)	Pre		23.8%	48.2%	24.9%	1.38	0.95	0.66	**0.69**	**0.48**
	Post		25.7%	44.8%	26.7%	1.33	0.90	0.53	**0.67**	**0.40**
Benchmark 6: % Suspect Population (All Crimes)	Pre		36.3%	36.3%	21.3%	0.90	1.26	0.77	**1.39**	**0.85**
	Post		32.6%	36.8%	22.8%	1.05	1.09	0.62	**1.04**	**0.59**
Benchmark 7: % Suspect Population (Part I Crimes)	Pre		37.1%	35.4%	21.3%	0.88	1.29	0.77	**1.46**	**0.87**
	Post		29.3%	38.7%	23.1%	1.17	1.04	0.61	**0.89**	**0.52**
Benchmark 8: % Suspect Population (Part I Violent Crime)	Pre		37.3%	35.4%	21.1%	0.88	1.29	0.78	**1.47**	**0.88**
	Post		23.4%	44.7%	25.7%	1.46	0.90	0.55	**0.61**	**0.38**

[57] The pre- and post-residential population percentages for Tables A.1 and A.2 are the same because all population-based benchmarks are derived from 2020 U.S. Census data.

Table A.1. Comparison of APD Arrest Racial/Ethnic Disparity Ratios Across Benchmarks Pre & Post-COVID

	Pre/Post COVID		Percent Race/Ethnicity			Disproportionality Indices			Disparity Ratios	
			White	Black	Hispanic	White	Black	Hispanic	Black	Hispanic
% Arrests	Pre	(N=31,515)	31.3% (9,863)	40.7% (12,815)	24.9% (7,835)	--	--	--	--	--
	Post	(N=13,439)	29.6% (3,975)	39.6% (5,322)	27.9% (3,744)	--	--	--	--	--
Benchmark 1: % Residential Population	Pre		43.5%	16.6%	29.0%	0.72	2.45	0.86	**3.41**	**1.19**
	Post		43.5%	16.6%	29.0%	0.68	2.39	0.96	**3.51**	**1.41**
Benchmark 2: % Suspect Population (All Crimes)	Pre		36.3%	36.3%	21.3%	0.86	1.12	1.17	**1.30**	**1.36**
	Post		32.6%	36.8%	22.8%	0.91	1.08	1.22	**1.19**	**1.35**
Benchmark 3: % Suspect Population (Part I Crimes)	Pre		37.1%	35.4%	21.3%	0.84	1.15	1.17	**1.36**	**1.39**
	Post		29.3%	38.7%	23.1%	1.01	1.02	1.21	**1.01**	**1.20**
Benchmark 4: % Suspect Population (Part I Violent Crime)	Pre		37.3%	35.4%	21.1%	0.84	1.15	1.18	**1.37**	**1.41**
	Post		23.4%	44.7%	25.7%	1.26	0.89	1.09	**0.70**	**0.86**