RACIAL DISPARITIES IN LAW ENFORCEMENT TRAFFIC STOPS: MEASUREMENT, INTERPRETATION, & INTERVENTION POSSIBILITIES

Michael Dolan Fliss

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology in the Gillings School of Global Public Health.

Chapel Hill
2019

Approved by:

Stephen W. Marshall

Frank Baumgartner

Paul Delamater

Charles Poole

Whitney Robinson

ABSTRACT

Michael Dolan Fliss: Racial Disparities in Law Enforcement Traffic Stops:
Measurement, Interpretation, & Intervention Possibilities
(Under direction of Stephen Marshall)

Law enforcement traffic stops are one of the most common entryways to the US justice system, with significant downstream impacts for both individuals and communities. Group-specific rates are typically based on jurisdiction resident populations; these rates, like many justice-system indicators, demonstrate race-ethnicity disparities. Residential-based rates implicitly assume race-ethnicity groups have equal vehicle access, equal driving volume, and that all driving occurs in resident's jurisdictions. In contrast, surveys suggest Black non-Hispanic and Hispanic households have less access and drive less than White non-Hispanic households.

Aim 1 reported the direction and degree of change in disparity indices when accounting for these driving factors. Data from over 20 million traffic stops in North Carolina were combined with US Census data and race-ethnicity driving factors from the 2017 National Household Travel Survey to calculate traffic stop rate-ratios (TSRRs) under multiple model assumptions. Spatial simulation models distributed Vehicle Miles Traveled (VMT) across the state and rebuilt rates for 177 law enforcement agencies. Adjusting for three driving factors simultaneously, disparity indices increased 15% on average from 2.02 (1.86, 2.18) to 2.33 (2.07, 2.59) for Black non-Hispanic drivers and were largely unchanged for Hispanic drivers. All models suggested both groups experience disparate traffic stop rates compared to White non-Hispanic drivers.

Aim 2 evaluated an intervention from 2013 to 2016 in Fayetteville, North Carolina that prioritized safety stops, intending to reduce both traffic crashes and disparities. Synthetic control methods were used to compare Fayetteville to a counter-factual Fayetteville that did not enact the intervention, built by the weighted combination of eight NC cities matched on pre-intervention measures (2002-2012). These models demonstrated reductions in crashes and disparities and, in contrast to the Ferguson Effect hypothesis, the de-prioritization of investigatory and economic stops was not associated with increases in crime.

Supplemental analyses explored the author's driving, alternate intervention evaluation methods, and within-jurisdiction spatial dynamics. The Public Health Critical Race Praxis (PHRCP) guided framing, results interpretation, and self-evaluation of the dissertation aims.

Traffic stops have associated public health outcomes and create disparities of relevance for public health researchers. Interventions guided by critical public health frameworks can save lives and reduce disparities.

To Tang Soo Do Mu Duk Kwan Master Ron Huntley,
the first Black man deeply in my life as a mentor.

Thank you for spending
a decade of your life
planting and nurturing seeds
of physical training, spiritual practice,
personal development, critical ethics,
and aspiring anti-racism
in the body-mind
of a young White child.

You changed my life.

ACKNOWLEDGEMENTS

After being introduced to Buddhist sutra and practice in my early teens and studying and sitting zazen ever since, and after my first anti-racism trainings in my early twenties, I have come to vigorously disbelieve what seems to be a very-human habit of mine: imagining that I ever did, and ever could ever, stand alone. If this is true of the parts of myself then it is doubly true for any of “my” work.

Instead, reflecting on this dissertation and on the upcoming competition of my doctorate in epidemiology, I am filled with appreciation for who and what has enabled me, and this work, to come this far. Since both I and this work are unfinished, I will carry this appreciation into everything that is next for as long as I exist to be and carry. In keeping with Buddhist traditions that center the three refuges, I offer this dissertation to the (1) Buddha, (2) Dharma, and (3) Sangha and their secular parallels: (1) the teachers and mentors who have instructed me, (2) their teachings and practices that have sustained me, and (3) my peer community of fellow activists, practitioners, and researchers who aim to apply that guidance and those teachings, encouraging each other along the way. I would be lost without these three gems in my life.

Buddha: my teachers and mentors

To my mother, who raised me and my brother with too little help, and who modeled hard work, emotional boundaries, and the value of education, who kept me safe as best she could: I love you Mom.

To Master Ron Huntley, my first martial arts instructor and surrogate father for a decade, who trained me physically, mentally, and ethically, who was my pivotal and first Black male role model, who pushed me as a child to grapple with ethics, who introduced me to meditation: you laid the foundation for the role of practices in my life, and these practices have kept me fed.

To Michael Gustafson, my signals and systems electrical engineering instructor and mentor during my undergraduate program at Duke, who helped me connect my disconnected intellectual enjoyment and passion for good: thank you for those conversations and late night chats; they were pivotal.

To Steve Kaufmann, my martial arts instructor and dear friend, who trained me for over fifteen years, who was patient with me as I grew to be a better student, teacher, and person, who continues to pass on lessons body-to-body and mind-to-mind, who introduced me to other teachers I never would have met without his lifetime of dedication to practice, who is the most sincere student in our school while also being its chief instructor: I'm so thankful for what you've given me.

To Calvin Allen, former executive director of Public Allies, NC, who gave me my first real job out of college, who introduced me to anti-racism work and the mission-driven world, who helped integrate my ideals of service with privilege responsibility and an intersectional worldview: you introduced me to deep activism in practice, and I'm forever grateful.

To Tema Okun, who built the first anti-racism models I'd ever seen, whose work on White supremacy culture in organization helped link my external community work to internal spiritual development, who always models humility and collective spirit though is one of the

most expert White people I know: your life's work didn't just help me in organizing, but laid the foundation for integrating my internal and external work as a White person.

To Scott Proescholdbell, who has championed me for many years, who continues to give me meaningful work projects I thrive on, who is a powerhouse for change and action at the state: you regularly inspire me. I am so looking forward to working more together after this.

To my dissertation committee, who encouraged me to take on a difficult, non-traditional project, who offered their timely feedback even when I have given then too-little time, who offered me time and space in their offices and classrooms, who emphasized that this dissertation is not my life's work, but a milestone in it: thanks for your patience, expertise, and encouragement. In particular, to Steve and Whitney, who took me on as advisees after Steve Wing's death: thank you for your compassionate support of many kinds.

To Steve Wing, who took me on as an advisee even after a rough start, who introduced me to the Environmental Justice Network, who modeled accountable activist scholarship, who first started this project with me and encouraged me to pursue it, who kept advising me as he was dying: I miss you terribly, and promise to work together your legacy of students and community partners to keep our work alive.

Beyond those I had a closer personal relationship with, I give thanks teachers from the lineages that Steve Kaufmann and I share (particularly Aikido, Chen Tai Chi Chu'an, and Japanese Zen) going back to their founders, passed down body-to-body and mind-to-mind. To those other teachers who I spent time with personally, some of whom are living and some now dead, including Yamada Shihan, Sugano Shihan, Chiba Shihan, Lehrman Shihan, Kongsburg Shihan, Grandmaster Wang Hai Jun, Richard Clarke, and Thich Nhat Hanh: the lineage you maintained and passed on enabled practice to arrive in me, and that practice has fed me for decades. I will not forget how training with or meeting you felt. Beyond those I have trained with, to the teachers I have never met, including the founders of my core practices and the authors of books on anti-racism: your practice has echoed for generations. Thank you for your gracious gifts as teachers. I stand small on your giant shoulders.

Dharma: their teachings and practices

To the study and practice of Buddhism, specifically Japanese Soto and Rinzai Zen, Zazen and Shikantaza, Vietnamese Zen in the Thich Nhat Hanh Lineage, Tibetan Gelugpa Mahayana Buddhism, the US American Buddhist traditions: You helped me turn the challenges of my teenage years into study and practice, and you have fed me ever since. I awakened to activism and my own privileges through bookcases of sutra and commentary and countless hours of sitting different forms of meditation alone and in groups. To the innumerable book authors recent and ancient, the professor who supported my Buddhism minor, to the local and distant zendos I have practiced at: just as emptiness is form and form emptiness, Buddhism must also be implemented to be experienced in the relative sense. I promise to never abandon the three gems or forget my time on the cushion.

To anti-racism education and practice, from James Baldwin to the People's Institute for Survival and Beyond: the world made so much less sense without a critical anti-racism, intersectional lens – especially as a White, able-bodied cis-man who passes for straight. I am in debt to the lineage of critical activists and scholars who have created, developed, and implemented those teachings in ways that matter for not just sense of self, but better ethical action in a racist, capitalist world. I promise to do what I can to better learn and pass on those teachings.

To my physical practices, particularly aikido and tai chi: I am so glad to have found physical arts that so directly link my healthy physical and mental development. I promise to continue to give you most of my evenings and many of my mornings.

To David Allen's Getting Things Done, which I have used as my life management system for nearly twenty years: I've been able to do more, with more accountability, thanks to this framework. The doing gives me joy, so I thank you for enabling me to juggle my diverse interests and commitments more effectively.

To epidemiology and public health, a career I've fallen in love with, with all its historical and present-day flaws: I still believe in this work, even as it depends on community organizing. May we improve together. I'll do my part.

And to all my practices, I am sorry to have relatively reduced my practice in the last few years, and months specifically, while juggling learning and work and activism. I'll be juggling those same things for as long as I can do anything, so I look forward to exploring new prioritizations after graduation. I promise to feed you as you've fed me.

Sangha: my community

To Margo, my partner of ten years, who encourages and inspires me, who shares so many values and practices: I love you. I look forward to what's next for both of us. You're an amazing person, and I feel so lucky to grow with you and benefit from your growth.

To my brother Tom, whose joy for life and learning are infectious, whom I believe in deeply, who grew up alongside me in martial arts: I love you. Here's to many years of conversations and outdoor activities. Perhaps one day I'll be reading your dissertation.

To my closest epidemiology friends, my brain trust, whom I met through Steve's advisee group, who held me through our sadness, who inspire me in your shared values, who are so creative and passionate and giving with your time, who have been co-authors and co-workers and co-conspirators on many projects: you were the most important part of my PhD experience. To Danielle, Libby, and Adrien in particular: I love working with you. Here's to a lifetime of creating trouble together.

To Pav, who's been my friend and sangha-mate for twenty years: I miss you! It's always so funny, and so expected, to share so many similarities in our paths, different as we are in some ways. Thanks for being with me for so many years in so many ways. I still think about your amazing PhD defense, and am so grateful to have you as a friend and scholar in my life. I look forward to attending your shodan exam!

To Carmen, who's been my friend for over fifteen years, now a fellow teacher, who's an inspiring academic and personal powerhouse, who's listened me discuss this dissertation work for years though in a completely different field: I'm so glad to still have you in my life. I aspire to your level of intuition and take-no-guff leadership. I finally finished my dissertation. You clearly beat me though.

To all my sangha, I hope to be able to spend more time and be more present with you after graduation. I miss you!

Offering

I offer this work as a White person aspiring to anti-racist theory and action. While aspiration is important, both Buddhist theory and anti-racism theory advocate for being accountable to our products of action, not aspirational intentions alone. I and this work are imperfect. Some of those imperfections I know of are conscious compromises for action in an imperfect world. Some are imperfections I simply acknowledge and regret. Other mistakes I do not yet have the clarity to see. As I will discuss later in the book, there are components of this work, its methods, and its implicit and explicit theoretical frameworks, that are lacking in anti-racism analysis and motivation. I hope this work does more good than harm, and though I have made efforts to try to ensure this, I cannot be sure.

In that vein, I humbly submit this dissertation as an offering to the Buddha (and my teachers and mentors), to the Dharma (and their teachings and practices), and to my Sangha (and community of friends, practitioners, activists, and researchers). I am here only because you are and have been here. May this work be of benefit.

To paraphrase the Bodhisattva Vow from Shantideva's Bodhisattvacharyavatara, which has served as my primary textual guide for now twenty-five years:

> For as long as space endures,
> For as long as beings are to be found,
> So too may the work for justice continue
> To dispel the sufferings and inequities of the world.

PREFACE

I was drawn into this work partly through happenstance. As a White (non-Hispanic) person, I'd been privileged to have disproportionately little interaction with law enforcement and experienced traffic stops only rarely. I was also privileged to be exposed to critical racism theory and anti-racism training explicitly in my early twenties – a rarer experience as a White person. Thanks to that training (and those mentors, see Acknowledgements), I was likely more aware than most White people of the history of racism in the United States and its consequent disparities in law enforcement outcomes for people of color, though that is a low bar. Through my work in public health, starting at the Orange County, NC Health Department (OCHD) in 2009, I'd had only a few professional interactions with law enforcement. I knew generally to be critical of law enforcement collaborations, knew well the role of policing in enforcing a racist, capitalist system (e.g. war on drugs, explicitly coordinating with KKK in some areas, surveillance of civil rights movements), but had done little direct work in the area. I'd had a long-time personal and professional interest in documenting and acting on disparities. But I did not expect to write an epidemiology dissertation on traffic stops when I started my PhD at the University of North Carolina, Chapel Hill.

In my application materials I shared my interests in public health informatics, particularly in indicator design and necessary infrastructure. Early on, after collaborating with my then-advisor Steve Wing on components of a Title VI complaint for siting hog farms in disproportionally Black North Carolina communities, I thought I might end up writing my dissertation on environmental racism. I didn't give much thought to traffic stops and their relation to public health.

That disconnect was part a function of my White privilege, certainly, but that privilege was reinforced by my experience in our school and department, both implicitly and explicitly. Implicitly, I remember exactly zero references to problematic partnerships with and resistance against law enforcement in core coursework, though a read of history reveals plenty to discuss. This absence reinforced my disconnect, leaving me unsurprised when fellow students and some professors would explicitly ask or directly state that policing had nothing to do with public health in general and epidemiology specifically. Though later in the program key faculty relationships did encourage me, my assumptions about the disconnect between public health and law enforcement were not challenged early on.

Instead, it was my relationship with my partner and my interest in local community action that drew me to the work. In my (now decade-long) relationship with Margo Krome-Lukens, a White woman, I am regularly appreciative of our mutual commitment to anti-racist action and regular discussion on each other's external and internal work (see Acknowledgements). Beyond those discussions and our own action in our professional capacities (she brings that critical anti-racism lens to her work in food systems), Margo was serving as the assistant secretary for our local NAACP chapter in 2014. I'd been a NAACP member for years and had been following NAACP associated initiatives but had not gotten involved in any committee work directly. I was also following the growing national conversation around police killings of Black individuals[1], many unarmed, and the associated increase of national press around issues of disparate policing.

In October of 2014, I had just begun my first semester of the Epidemiology PhD program, and was still concurrently working as an epidemiologist with the Orange County Health Department, paying my way through school instead of obtaining a traditional graduate research assistantship. Locally, I had worked briefly with Chief Chris Blue of Chapel Hill Police Department in my epidemiology and informatics role at OCHD, as part of a overdose collaborative the lead Orange County to a number of firsts in the state around opioid-epidemic-related interventions (e.g. naloxone carrying by officers, distribution in health departments); however, that health department collaboration was not the bridge to this work. Instead, I came across an article citing work in Durham by the Southern Coalition for Social Justice (whom I would eventually collaborate with in the coming years). I did not yet know that the early work of Frank Baumgartner here at UNC Chapel Hill was also informing much of this national press. I reached out to Margo, wearing my professional hat as a math-savvy health department employee, in her role with the NAACP. Citing that article, I offered my help to our local NAACP branch if crunching data were useful for conversations[36].

In December of 2014, Margo and Barbara M. Foushee, branch Secretary, reached out to membership to notify member of an upcoming opportunity to engage Chief Blue of Chapel Hill Police Department, and Chief Horton of Carrboro Police Department, and Orange County Sheriff Blackwood. I submitted these questions in advance:

[1] Death by police killing is formally known by the phrase "death by legal intervention." This phrase is both sanitized and problematic. There has been recent literature on the public health responsibility to investigate and reduce these deaths. Alongside this dissertation I also have a paper, co-authored with another student, on the use and limitations of the National Violent Death Reporting System to better count these deaths, as called for by Krieger et al. in 2015[86].

- Does your staff regularly review traffic stop data in aggregate, looking at race distribution and stops in low-income areas? If so, what does your staff make of the findings? If not, why not?
- In both Chapel Hill and Carrboro, Black folks made up around 25% of the stops (but are only around 10% of the population). Why is this?
- In many categories of stops and overall, Black drivers were 2-3 times as likely to be searched as white drivers. Why is that?
- The overall racial difference in searches after stops seems to be increasing over time. Why is that the case?
- Unlike seatbelt, speed limit or stop light violations, some of the categories of stops and searches seemed particularly subjective, like "investigation" or "some other motor vehicle violation". These reasons showed some of the highest racial differences in both Chapel Hill and Carrboro for Black and Hispanic drivers. Why is this?
- Some of the causes for stops (equipment or regulatory lapses) would be expected to more associate with those living check-to-check. This would increase the number of stops for those with low incomes, who have more difficulty in keeping up their insurance or vehicle repair. What are your departments doing to combat this increase in stops? What policies have you heard of that could assist in easing this burden on low-income people?
- Overall, what do you think of a situation where Black folks are being both stopped at a higher rate than white folks when compared with the population and then searched at a higher rate after having been stopped? What is the intended effect of this

difference in enforcement, even if unofficial and unintentional? And as a follow-up: Is this difference in enforcement intentional?

Wearing my official local health department hat for the data request, for the first time I reached out to the NC State Bureau of Investigations, the data owner of the North Carolina traffic stop dataset. They pointed me to aggregate data at the NC SBI website, but also offered to send me a CD of the dataset that would eventually serve as the basis for this dissertation. Using that data I built the following table (see below), the earliest deliverable from this project, in December 2014 in advance of the community meeting.

This earliest work is not without significant flaws, perhaps most notably (1) my failure to create a combined race-ethnicity variable, leaving Hispanics in the White racial category to which reduces the disparity of Black non-Hispanic people toward null, and only a hand-waving note about residential and driving dynamics. However, feedback from the branch made it clear these numbers, in a readable format, were very useful with police chiefs and the Sheriff. Most directly, when the Orange County Sheriff said publicly that his goal was to have half white and half black traffic stops, a member of the audience used this simple table to reply that Blacks made up only 12% of the population of the county. This was seemingly news to the Sheriff, challenging his implicit benchmark for agency equity, and this quote made it into a local newspaper article about the event. This clarified for me three things: (1) traffic stop data was widely being interpreted without regard to underlying rate dynamics, (2) because of this, law enforcement (at least) were not considering differences in demographic representation when considering traffic stop disparities. I naively thought a little work could go a long way in assisting those two issues.

Chapel Hill	White	Black	Native American	Asian	Other	Total By Race	Hispanic	Non Hispanic	Total By Ethnicity
Total Stopped	13554	5182	177	1248	102	20263	1470	18793	20263
Total Searched	314	295	2	15	1	627	61	566	609
% Searched of Stopped	**2.3%**	**5.7%**	**(S<10)**	**1.2%**	**(S<10)**	**3.1%**	**4.1%**	**3.0%**	**3.0%**
% of all stops	67%	26%	1%	6%	1%		7%	93%	
% of CH population	73.2%	8.6%	0.4%	13.5%	4.3%		5.2%	94.8%	

Carrboro	White	Black	Native American	Asian	Other	Total By Race	Hispanic	Non Hispanic	Total By Ethnicity
Total Stopped	8965	2612	62	561	58	12258	1348	10910	12258
Total Searched	236	192	1	7	1	437	68	369	428
% Searched of Stopped	**2.6%**	**7.4%**	**(S<10)**	**(S<10)**	**(S<10)**	**3.6%**	**5.0%**	**3.4%**	**3.5%**
% of all stops	73%	21%	1%	5%	0%		11%	89%	
% of Carrboro pop	70.8%	7.9%	0.1%	8.4%	12.7%		16.3%	83.7%	

Sheriff	White	Black	Native American	Asian	Other	Total By Race	Hispanic	Non Hispanic	Total By Ethnicity
Total Stopped	2287	963	12	58	166	3486	452	3034	3486
Total Searched	94	53	0	3	19	169	45	124	147
% Searched of Stopped	**4.1%**	**5.5%**	**(S<10)**	**(S<10)**	**11.4%**	**4.8%**	**10.0%**	**4.1%**	**4.2%**
% of all stops	66%	28%	0%	2%	5%		13%	87%	
% of Orange County	75.5%	11.5%	0.4%	7.2%	2.4%		8.2%	91.8%	
% of North Carolina	69.8%	21.5%	1.2%	2.3%	5.3%		8.7%	91.3%	

Notes: Total stopped by driver race. Total searched is only of drivers (ignoring passenger searches). We are not calculating population-based rates because of factors that would confuse these rates, like different vehicle ownership, cross-country or policing agency travel, and commuting patterns - these are not necessarily stops of Chapel Hill and Carrboro residents, but stops from these agencies (which likely includes a large number of residents). However, it is important to remember the particular racial and ethnic demographic distribution when comparing these crude counts for stops. Demographic data for the closest geographic unit to the precinct from the 2011-13 3 year ACS estimates is offered to facilitate conversation. Stop data from NC DOJ SBI webpage for Jan 2012-Oct 2014 (Nov and Dec data unavailable; http://trafficstops.ncdoj.gov/).

Figure. Earliest table prototype from December 2014.

While I had been following recent articles on racial disparities in traffic stops, I had not yet realized one of the main authors of cited studies on traffic stop racial disparities was at UNC Chapel Hill (Frank Baumgartner). While searching for data and articles on traffic stops, I happened to come across Frank’s name and website, where he had set up a dedicated page for his traffic stop projects[2]. I read all his early whitepapers, then finally reached out in late December to introduce myself and set up a first meeting. In those white papers he and his student collaborators often focused on searches, and often using odds ratios. While appropriate for his aims, I also knew that odds ratios were notoriously difficult to interpret for most non-scientists (save, perhaps, gamblers). For this reason, issues of rate building factored heavily in my first email on the subject.

I was able to be an early collaborator with Frank on what seemed to be odds-and-ends, albeit important ones, of his continuing analyses: some spatial analysis and map building, discussion of the consequences of separating race from ethnicity in analysis, and feedback on a few chapters of what turned into his now-published book. Frank's work on traffic stops is far reaching, and with dedicated graduate and undergraduate student collaborators in his department, he was able to analyze many questions relative to the distribution of searches, the role of outlier officers in driving agency disparity metrics, associations with political power, and more. The more we discussed, the more we came to leave to me the underlying issue of improved stop rate denominators. Reflecting now, I remember distinctly wishing my analysis could both move faster and be more broadly applicable as he was preparing his authoritative book on NC traffic stops. Given those conversations began now four and a half years ago, and given what strike me as modest aims, I am struck by how much work there is still to do; work that feels like should have been completed years ago (see Next Steps in Discussion).

For most of the first year of my work on this project, I assumed this would be a community project that would produce no scholarship-related deliverables. Coming from public

[2] https://fbaum.unc.edu/traffic.htm

health practice at the local and state level, I know well that a large amount of public health action (not to mention community activism) does not produce peer-reviewed dissertations and published journal articles. Regardless, my then-advisor Steve Wing gave me early feedback on this project that first and second semester, even before the possibility of it having any academic connection. Also, in that first year, I had early conversations with Charlie Poole, who graciously donated hours of his time for discussion of this community project in his office. Early conversations with Charlie were around the primacy of measures on the additive scale vs. multiplicative scale for communication – these issues remain important to this project.

As discussed previously, I remember no discussion of policing during our core epidemiology methods sequence. However, and to her credit, Julie Daniels supported me during our required grant-writing class in my second year by her welcome allowance of my non-traditional project. Only one of the aims proposed in that grant-writing project is represented here in this dissertation (Aim 1 on improved measurement); the other aims are now relegated to next steps (see Discussion). However, Julie also directly spoke up in my defense when some students suggested I drop the project entirely because of its irrelevance to public health. At the time the project had no direct injury component (Aim 2 on the Fayetteville intervention was not yet conceptualized) and was only focused on accurate measurement of stop rates. I am indebted to her for the space she gave to develop these ideas and provide early feedback.

Beyond the core courses, I'd be remiss if I didn't mention two places law enforcement came up in my subject-area tracks: (1) in Whitney Robinson's social epidemiology course, where Frank Baumgartner (a political science professor, to be clear) guest lectured on his extensive work on racial disparities in application of the death penalty and (2) in Steve Marshall's injury epidemiology sequence, where some law enforcement related injury collaborations were discussed. Again, I am thankful to have their feedback reflected in this dissertation, even though I presented class projects before I was sure they were dissertation material.

Through Frank I was connected to Ian Mance at the Southern Coalition for Social Justice (SCSJ). SCSJ was my only funder of this work, with a one-time stipend for me to write up some of my early findings in a white paper summarizing Fayetteville's intervention. Much of that white paper was on the sub-agency dynamics in Fayetteville, a previous aim that has since been dropped from this dissertation. Some of that work is retained in an Appendix as further evidence of Fayetteville Police Department's enacting the intervention. Ian is the lead on the Open Data Policing website, a new resource that makes the traffic stop data from the difficult to use NC SBI website available to the public. That website has grown to include data from multiple states. I have been able to serve as a technical advisor on some of the visualizations and underlying data processing (such as the handling of race-ethnicity variables from the American Communities Survey).

Through Ian I met now retired, then Fayetteville Police Department Chief Harold Medlock. He conveyed the real-world challenges of policy implementation within a large police department. Those conversations were invaluable to understanding important nuances in the intervention, even if they cannot all be quantified in numbers. Beyond those conversations we shared, he has been a tireless advocate for a more public health-oriented policing, speaking as a champion to other agencies. I knew of his leadership around the overdose epidemics but had never had a chance to discuss policing with him personally. I am indebted to him for his time and advocacy from within.

Through Ian and Frank I was connected with the Orange County Bias-Free Policing Taskforce, a NAACP affiliated workgroup, in 2015. James Williams, Jesse Gibson, Rich Rosen, and Tye Hunter are the core of this group of lawyers, professors, community organizers advocating for policing change at the law enforcement agencies that patrol Orange County. There we've drafted statements on policies and tracked many measures of policing, including but not limited to traffic stops, and I've attended and spoken at meetings on the subject. I am truly grateful to have had the chance to serve on that taskforce over these recent years, and only regret my research didn't progress faster and I have been less available in the last few months as I've worked to finish this dissertation. This local experience meeting with police chiefs and fielding questions about this data in a real way has been a boon to this project. I am also appreciative of the time of those local police chiefs and sheriffs.

As this preface documents, this dissertation is more than two publishable papers and some contextualizing chapters. It is instead just a small part of a much wider organizing effort for police accountability and equity. It is neither the pinnacle of analysis of disparities, nor the most progressive framework for rethinking public safety and enforcement. Enough organizers and researchers suggested that it's two modest aims would be useful steppingstones to enter into the peer reviewed literature. We seek to (1) provide evidence to suggest integrating travel realities is essential when considering traffic stop disparities, even if they are already large, and (2) document a novel intervention and the underlying thinking behind it. These steppingstones are best seen as an attempt at harm reduction than any sort of solution or ideal, and carry with them all limitations associated with incremental change. There is more to say and do about traffic stops and their apparent disparities. I expect to continue to follow community organizing, learn from experts of all kinds, and contribute my efforts toward equity where useful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACS	American Communities Survey
API	Application Programming Interface
CRT	Critical Race Theory
DiD	Difference-in-Difference
DMV	Division of Motor Vehicles
FPD	Fayetteville Police Department
LEA	Law enforcement agency, e.g. police department, sheriff department, etc.
LEHD	Longitudinal Employer-Household Dynamics
LODES	LEHD Origin-Destination Employment Statistics
LATCH	Local Area Transportation Characteristics for Households
NAACP	National Association for the Advancement of Colored People
NHTS	National Household Travel Survey
NC	North Carolina
PHCRP	Public Health Critical Race Praxis
SCSJ	Southern Coalition for Social Justice
SBI	State Bureau of Investigations
TSRR	Traffic Stop Rate Ratio
US	United States

CHAPTER 1 - INTRODUCTION

1.1 Background

Growing out of a history of explicit discrimination, Black and Hispanic individuals remain both overrepresented in and disproportionately impacted by the US justice system [118]. Disparate justice-related outcomes, including court and fine practices [132], the application of the death penalty [21], and use of excessive force [100,122] have severe economic and health impacts for individuals involved and their communities. National press [120] and community groups [98] have highlighted these disparities after videos of unarmed people of color being shot by police were released publicly, often during or following a traffic stop [8]. Community-led movements [6], national press [120], peer-reviewed research [16] and the Department of Justice [132] have all suggested that traffic stops are most burdensome to low-income drivers and their communities and are a significant indicator of systemic race- and income-based discrimination. Law enforcement traffic stops are one of the most common entryways to the US justice system, with significant downstream impacts for individuals and communities. The limited data we do have suggests 12% of all drivers, and twice as many racial minority drivers, are pulled over each year by law enforcement [39]. Yet states have only recently required agencies to collect and report these stops.

As important as traffic stops disparities may be as an indicator of disparities and as a primary entryway to the justice system, the technique for estimating these police traffic stop rates is known to be fundamentally flawed [132,145]. Known in criminal justice literature as “benchmarking”, this technique builds traffic stop rates based on **residential populations** denominators of police county or city jurisdictions as a proxy for either their driving populations or vehicle miles driven at risk for stop [48]. However, assessing race disparities in stops requires estimating race-specific rates based on the **driving population truly at risk of stop**. Preliminary national data on race-specific driving patterns, like differences in vehicle ownership by race (e.g. 51% of Black households have vehicles vs. 84% of white households [91]) suggest the already disparate rates by race based on residential populations **may widely underestimate the true disparities.** Because policing practices, populations at risk, and rural-to-city driving flows can vary widely between city and county jurisdictions, and because policy change often happens at the specific jurisdiction level, communities and law enforcement agencies (LEAs) require these stop rates to be jurisdiction-specific.

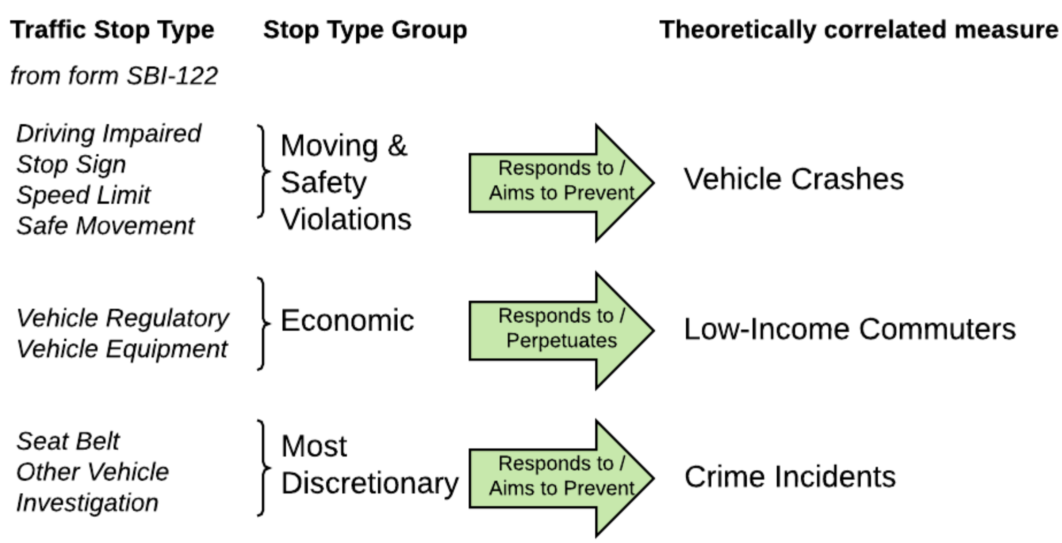


Figure 1.1. Theoretical stop type intended targets of intervention.

Besides considering agency-specific and travel-informed stop rates by race-ethnicity, not all traffic stops are the same. North Carolina's traffic stop database, one of the oldest and most complete in the nation [16], captures ten kinds of traffic stops. For the purpose of this discussion, we divide traffic stops into three categories: (1) “safety stops” including violations of speed limits, stop lights, driving while impaired, and safe movement; (2) “investigatory stops” including explicit investigation, unspecified rationales, and discretionary seatbelt enforcement; and (3) “economic stops” including vehicle regulatory and equipment violations, such as driving without a license, insurance, registration, or a completely working vehicle.

Relatedly, it is important to note that LEAs do not operate in a vacuum, and that public health can influence and inform LEA priorities. As an example, here in NC the “Click-It or Ticket” program, conceived in partnership between the UNC Highway Safety Research Center and NC Department of Health and Human Services, has enlisted local LEAs to prioritize seatbelt stops in some jurisdictions, with the intended goal of reducing severity of and fatalities from traffic crashes. However, it is important to consider the disparate impact and implementation practices of seemingly group-agnostic public health interventions, especially to marginalized populations. As an example from a different body of literature, studies of a tax on non-essential foods in Mexico show a disproportionate impact on food purchases by socioeconomic status [14,29].

Further, a **system dynamics perspective** suggests the importance of considering downstream negative effects, hidden feedback loops, and collateral harms of interventions and in modeling. In this case, preliminary findings show that of the ten stop-type reasons, seatbelt stops have one of the largest apparent racial disparities – suggesting they may also be used as an excuse for individual race- or neighborhood-specific pretextual stops. But public health focused on traffic safety, injury prevention, and harm reduction has other intervention tools available that may not disproportionally impact low-income and marginalized race-ethnic communities in the same way as seatbelt stops may. Though this is just one example of the intersection of public health practice and policing priorities, it is clear the measurement, evaluation and interpretation of stop disparities must be done in the context of larger public health interventions that address similar end goals but have different collateral impacts and equity considerations.

	Raleigh Population, '15		% Black/AA
	Total Population '15	*439,896*	*29%*

	Raleigh Traffic Stops, '02-'13		% Black/AA
Moving & safety violations	Driving Impaired	10,025	26%
	Stop Sign	46,609	37%
	Speed Limit	216,451	37%
	Safe Movement	39,924	41%
Economic	Vehicle Regulatory	215,598	49%
	Vehicle Equipment	74,500	55%
Most discretionary	Seat Belt	23,529	46%
	Other Vehicle	60,598	49%
	Investigation	32,481	54%
	Total Stops '02-'13	719,715	44%

Table 1.1 Suggestive racial disparities in traffic stops by stop reason, Raleigh, NC.

Agencies and officers have wide discretion in the application of traffic stops and prioritization of stop types. Similar to stop and frisk programs [89,90], court cases have been central in establishing the legal rationale for this discretion. Supreme court cases in 1968 and 1996 [26,80] enabled US law enforcement, under any reasonable suspicion and the loosest definitions of crime profiles, to escalate any traffic violation, however minor, into a traffic stop [16]. When combined with the driving reality that nearly all driving trips include actions interpretable as infractions, whether small wavering within lanes or movement over or under posted speed limits [16,89], these rulings permit law enforcement nearly complete discretion over traffic stop enforcement legally, even if the public views those stops as unfair [90].

Preliminary estimates show that race-ethnic disparities can vary widely by stop type, with the most subjective investigatory stops having the highest disparities and safety-related stops having the smallest. Recognizing these disparities, given the aforementioned discretion, and with a history of public health collaborations (e.g. in overdose prevention), the Fayetteville, North Carolina Police Department enacted an intervention in 2014 designed to save lives and reduce racial disparities in police stops by prioritizing safety stops significantly above others, moving from 30% safety related stops to 90% safety stops over a three-year period. Unlike most police departments, Fayetteville began geocoding its stops at the point level in 2013, allowing for a neighborhood-specific evaluation of the intervention's effectiveness of redeployment strategies and its impact on reducing racial disparities and injuries. As law enforcement agencies increasingly geocode their activities, this evaluation can inform efforts to reduce racial disparities in stops and promote injury prevention efforts through spatial targeting and explicit prioritization of preventable injuries (See Supplemental Analyses).

In 2013, Chief Medlock (retired in 2017) of Fayetteville Police Department (FPD), in part due to community pressure, directed officers in his traffic stop program to significantly reprioritize safety-related traffic stops over economic and discretionary stops, with the intended goal of reducing traffic fatalities and possible side-benefit of reducing disparities in traffic stops. As part of that effort, Fayetteville elected to begin to collect GPS point locations of traffic stops - one of few LEAs to do so in the state at the time, though given the rapid increase in availability of low-cost GPS tools more agencies should have access to point-level traffic data in the future. Per discussions with the Chief, the implementation was difficult, with some officers leaving or let go because of a difference in policing philosophy. However, after two years of these efforts (see below), in 2015 Fayetteville celebrated a reduction in traffic stops, a significant uptick in the percent of those stops that were safety-based and a reduction in traffic fatalities while the state saw an increase in the same. Fayetteville collaborated with the Southern Coalition for Social Justice (SCSJ), a legal non-profit in Durham that set up a dedicated website to increase transparency on traffic stops (www.opendatapolicing.com/nc/), to help monitor those efforts – The author created a preliminary white paper on these efforts for Fayetteville and served as technical assistant for the website.

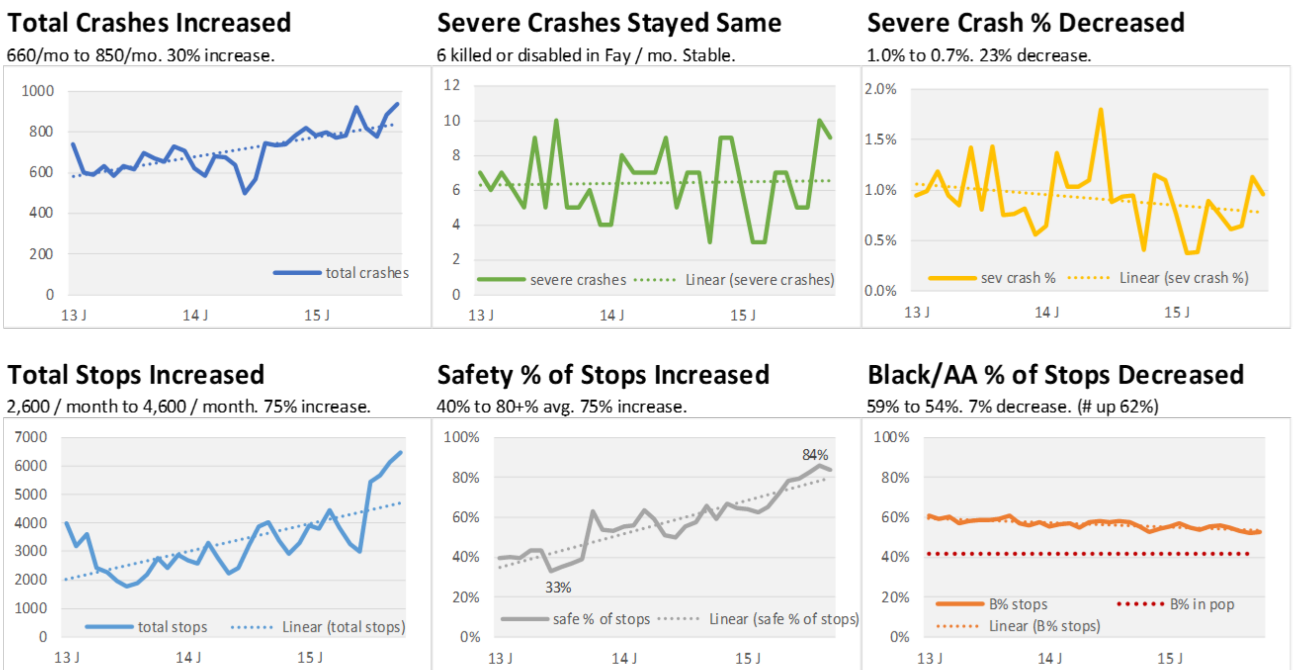


Figure 1.2 Traffic-stop related measures during Fayetteville intervention.

1.2 **Strengths**

North Carolina (NC) is uniquely situated to answer traffic stop research needs, with the nation's most complete and long-running statewide electronic police stop database of over 18 million reported traffic stops, publicly available, after requiring reporting on a consistent form by state statute in 2002. Though a near-census of traffic stops across the state, this SBI-122 data capture form (see Appendix 1) has two important limitations that contribute to the need for this research. First, that form does not capture global position system data (GPS, e.g. latitude and longitude) of individual traffic stops, only that a traffic stop was made by a given LEO in a particular LEA. Though a few LEAs elect to supplement this form with point-location data of the stop (e.g. Fayetteville PD, enabling aims 3 and 4), this disallows small-area analysis of traffic stop patterns within or between jurisdictions. Secondly, SBI-122 only captures the city and county of stop, not the residential city and county of the stopped driver or passengers. This increases the difficulty of understanding the underlying driving patterns and population at-risk of stop within a LEA's jurisdiction. While other traffic stop datasets may retain this information on driver residence, either by additional fields on the form or linking to license and registration information, adding this field by itself does not solve the underlying problem of appropriate at-risk driving populations. Neither rates built from residential populations of the stopping jurisdiction or of the residential home jurisdiction appropriately model traffic stop patterns; only driving-informed denominators accurately model the at-risk population for traffic stop rates.

This research has direct and actionable policy implications. Even when using flawed residential populations to approximate true stop rates and build race-specific traffic stop incident rate ratios (TSRRs), some police agencies have formally shifted policy by choice or public pressure. Recognizing apparent race/ethnicity disparities in stop reasons and search rates has lead agencies to enact policies including: mandating written consent before search; formal non-discrimination and prohibition of racial profiling policies; mandatory racial disparity training for officers; including stop rate data in officer review; and setting marijuana use as lowest enforcement priority. Particularly because of the immediate policy consequences to disparities this severe, accurate assessment of stop rates using at-risk drivers and vehicle miles driven, not resident populations (Aim 1) and understanding relationships to related ecological variables like crime, injuries, and poverty (Aim 2) are both essential and timely.

The White House administration has formally advocated for increased use and open sharing of policing data, launching the Task Force on 21st Century Policing in December of 2014 by Executive Order [23]. This initiative which will both further incentivize states to develop and use traffic stop indicators and push for novel utilization of increasingly detailed stop data. Following White House recognition, the Southern Coalition for Social Justice (SCSJ), a NC-based law and advocacy nonprofit part of the UNC-CH research coalition, had been tasked with supporting agencies in North Carolina to further open policing ideals described in the Executive Order. SCSJ has worked with UNC-CH researchers to build the nation's first open policing website (https://opendatapolicingnc.com/), allowing citizens, police chiefs and judges to search for agency-wide and officer-specific policing patterns in the entire NC stop dataset. It will be updated continually going forward and represents over 95% of the state population by police jurisdiction. This study is based on the same dataset as is currently in use on the open data policing website, and plans exist to grow the website as this research informs interpretation of these key stop variables, allowing lessons learned from this study to be automatically carried forward in time. Other states have begun to reach out to SCSJ to investigate building similar websites for open police data. Lessons learned in this research will extend out through these channels.

This research also anticipates future data needs and public health initiatives. First, more and more states are bringing similar, state-wide police traffic stop databases online, tracking police stop variables (like demographics, location, reason for stop) and consequences (search, arrest, use-of-force). States without current databases, like California, have passed legislation to mandate state-level reporting of traffic stops, to begin in the next few years. These states will benefit from precedent and insight set by this research. It is reasonable to predict that as technology rapidly increases, more police departments will be expected to maintain traffic stop data and utilize it meaningfully for evaluation and timely public health action. As a first example, collaborations between public health and police departments in response to the prescription drug epidemic have led to many officers newly carrying the opiate-overdose reversing drug Naloxone (Narcan). Because police are often first responders before EMT staff, North Carolina has seen dozens of lives saved by this initiative. Because time in an overdose is of essence, stop databases that can accurately describe patrol priorities can be used to focus patrol in areas of high overdose. As a second example, geocoded traffic and pedestrian/bicyclist injury data can increasingly inform police patrolling, and increasingly accurate databases of patrolling patterns will be able to be used to inform and evaluate patrol decisions for injury prevention. This UNC-CH research team has begun working with one such agency, the Fayetteville police department, who believe their conscious reallocating of police presence has not only reduced racial disparities in traffic stops without negatively impacting crime rates, but also has significantly reduced their city-wide traffic fatalities by patrolling high injury intersections. Each of the questions, tying public health outcomes to police activities, require accurately describing police stop rates so that they could be used in a model. The current residential-based rates are insufficient for exploring these future questions.

1.3 **Significance**

Preliminary research with the NC traffic stop dataset suggests widespread, significant, and disparate impact to marginalized populations (Baumgartner et al., 2018). Black and African American people make up 22% of the NC resident population but 31% of the 18 million stops analyzed in this dataset [Table 3, Preliminary Data]. Vehicle searches may follow stops, and preliminary analysis of the NC police stop dataset suggests that while Blacks and Hispanics are searched more often in most jurisdictions [Figure 1, right], police find contraband at similar or lower rates than searched white drivers (also see Baumgartner, 2018). Racial disparities in stops that lead to consequent disparities in searches may contribute to racial differences in arrests for drug possession without intent to sell, which have high racial disparities though surveys suggest drug use is generally similar across race [27]. Lastly, and most extreme, many recent deaths by police or while in police custody have occurred following stops for these economic or discretionary stop reasons [8,139]. These downstream consequences are varied and severe, whether financial, legal, imprisonment or loss of life.

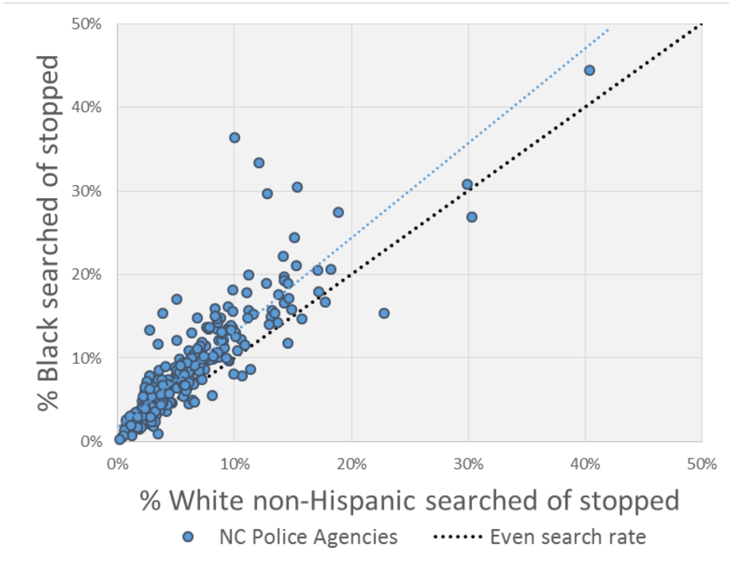


Figure 1.3 Racial disparity in searches after stop, NC police agencies.

Preliminary data is available for agency-level crude resident population rate, and simple adjustments (based on ACS and NHTS 2009, see section C, Research Design & Methods, for details) are presented below. Noting the many significant limitations with using resident populations to create stop rates (the substance of this dissertation), crude resident population rates suggest nearly all police agencies stop Black/AA drivers (and Hispanic drivers) significantly more often than White Non-Hispanic drivers (also see Figure 6 in following pages for Black/White Incident Rate Difference distribution).

Further, as described in the background and evidenced in Raleigh, NC (see table, previously), in many agencies this apparent racial disparity increases for stop reasons most connected to social justice priorities, especially those connected to income ("vehicle regulatory" and "equipment" stops) and the most discretionary or subjective stops ("investigation", "other vehicle" and "seat belt"). Preliminary data suggests the proportion of Black drivers stopped may double for these stop reasons in many jurisdictions, such as Raleigh, NC, where Black/AA drivers make up around 50% of stops for these justifications but 29% of the resident population. Though the residential denominators limitations are the central subject of this study, driving adjustments may suggest disparities are higher than this. Community groups report similar experiences of disparate policing nationwide; if these crude effects are similar nationally, the total disproportionate stop burden is immense.

While stops have consequences by themselves, stops also have significant downstream social and economic consequences for drivers, especially low-income individuals who are disproportionately People of Color. Stops with citations can carry steep fines difficult for low-income individuals to pay. These fines can force residents to choose between rent, food and healthcare, and contribute to a cycle of indebtedness and criminalization that, in its extreme, the Department of Justice characterized as predatory [132]. Even with clear, crime-related intent, frequent LEA stops can create experiences of discrimination and mistrust that permeate communities [56], can have concrete health effects, such as a chilling effect on 911 calls [65], may not prevent crime [9], and may constitute a human rights violation with significant collateral impact on those who have not violated any crime (e.g. stopping 100 Hispanic drivers with little cause aiming to catch one who may be engaged in illegal activity) [18]. Accurately measuring agency-specific stop rates (Aim 1) is essential to understanding the true impact of these policing patterns.

Though resident population-based comparisons are in widespread use and suggest significant racial disparities may be present, known confounders suggest these stop disparities may be even more extreme. First, county agencies often count entire county residential populations, even though they may functionally only patrol rural areas outside of cities that have their own dedicated municipal police departments. Given urban/rural populations are often different by race, assigning only rural populations patrolled by county departments may significantly change their racial demographics. Further, the US census suggests vehicle ownership by household is lower and use of public transportation is higher for People of Color (e.g. 51% of Black/AA and 84% of White non-Hispanic households have access to vehicles), suggesting even adjusted resident populations significantly underestimate this effect [91]. Simple adjustment by these two measures, using block-level census data by race to assign functionally policed residential populations to jurisdictions and using these national estimates of driving populations to recalculate IRRs, increases Black/White police stop incident traffic stop rate ratios (TSRRs) by 100% on average (mean +17%, max 550%), effectively doubling the disparity in stops we observe. As example (below), adjustment of Wake County Sheriff Department's TSRR denominators from the residential population to (1) the rural population it functionally polices then (2) estimating its driving population, raises its Black-white stop IRR disparity from 1.4 to 2.5 to 4.2 with both adjustments. This study proposes a more nuanced approach to driving adjustment, but even simple adjustments demonstrate the degree of potential bias in the current estimates, the need for true driving-based estimates, and the role of a sensitivity analysis in determining the possible direction of those biases.

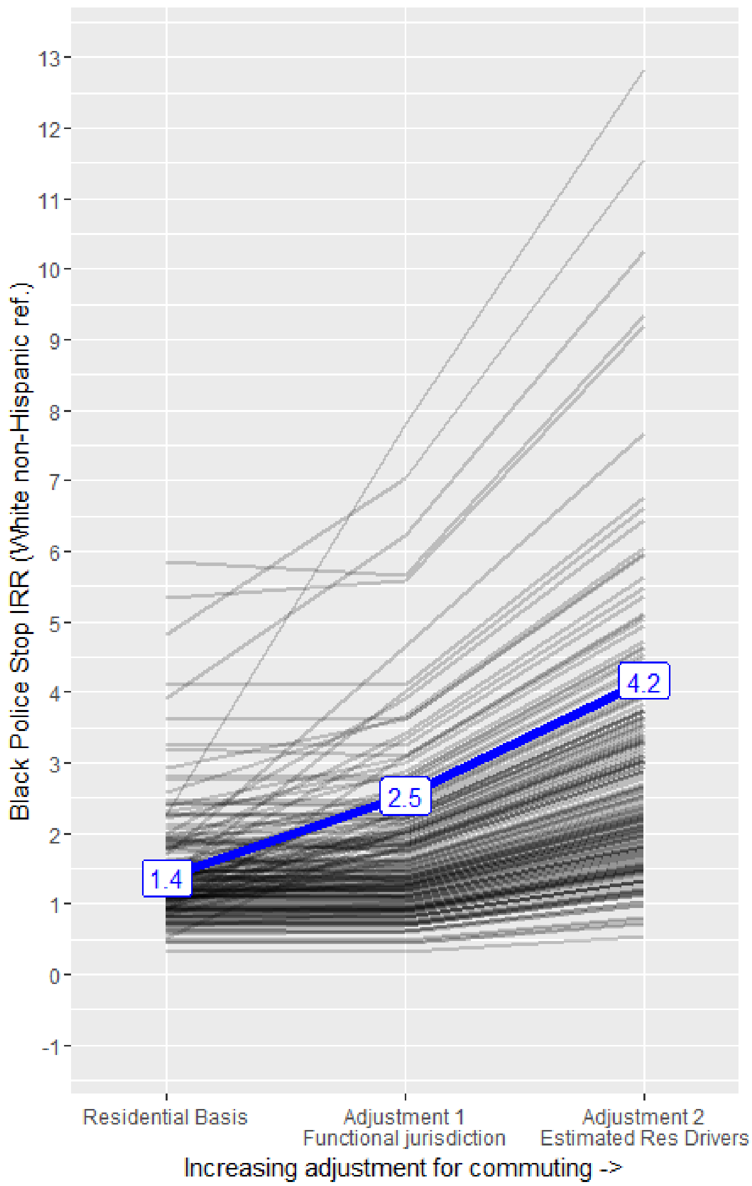


Figure 1.4. Demonstration of driving adjustment increasing racial disparities in stop rates.

On average police agency TSRR Black/White IRR disparity ratio increase by 100% with simple adjustments. As example, adjustment of Wake County Sheriff Department's IRR denominators from the residential population to (1) the rural population it functionally polices then (2) estimating its driving population, raises its Black-white stop IRR disparity from 1.4 to 2.5 to 4.2 with both adjustments. Includes 189 NC police agencies with more than 1000 recorded stops who had searched both Black and White non-Hispanic drivers. Analysis by Michael Dolan Fliss, UNC-CH, using NC SBI Stop Data, 2002-2013.

This research is timely, with immediate application nationwide. Baumgartner et al. [16]have worked for years documenting disparities in traffic stops and consequent searches, moving the literature on racial disparities forward. However, there have been no statewide analyses of traffic stop data that adjust for driving realities and considers directly public health focused ecological variables when interpreting those rates, even though literature consistently acknowledges these limitations. The few states with the beginnings of statewide police traffic stop databases have yet to do statewide analyses. Community and journal research on traffic stop rates has focused on individual jurisdictions and ignored questions of transport between jurisdictions in calculating stop rates [25,94,145].

1.4 **Conclusion**

This project ties together LEA stop research needs at multiple levels under a coherent public health framework, offering improving research techniques and frameworks for surveillance, public messaging, community oversight, and intervention design. The sensitivity analysis and consequent improved denominators and rates (Aim 1) can help inform improved surveillance of disparities in traffic stops by both individual law enforcement agencies, state and local oversight entities, and interested community groups. Fayetteville offers an opportunity to demonstrate public health impacts of efforts to curb disparities and reduce traffic harms (Aim 2). Following the literature review, the methods chapter and appendices include supplemental analyses exploring small-area modeling techniques and additional chapters demonstrate application of critical frameworks to understanding traffic stop programs. The discussion chapter enumerated strengths and limitations, re-applied critical frameworks as dissertation self-assessment, and describes areas for both future academic research and anti-racist action.

CHAPTER 2 - LITERATURE REVIEW

A review of literature searching PubMed and Google Scholar was completed using the phrase “traffic stop,” with supplementary keywords including “measurement,” “disparities,” and “bias.” with a particular focus on criminal justice and public health journals. Because of the relative novelty of traffic stop databases, white papers, law enforcement, and government reports were also reviewed using Google searches for the same terms.

Issues of measurement, framing, and action on traffic stop disparities are current organizing focus. These lessons may not have made it into formal reports or peer-reviewed literature yet. Therefore, this formal literature review also benefitted from active community collaborations and organizing around these issues.

2.1 **Traffic stops, disparities & their measurement**

A thorough history of the origin of policing and traffic stops is well beyond the scope of this dissertation and has been covered in popular press in recent years [2]. However, some historical context is important to understand the origin and growth of traffic stops as a law enforcement intervention.

In very brief, law enforcement traffic stops began soon after the introduction of motor vehicles, a disruptive technology on roads previously occupied by walkers and horse drawn carriages [119], in the later 1800s. Skipping ahead fifty years, The Green Book [59], a travel guide for Black motorists published from 1936-1967 by Victor Green during the Jim Crow era, describes areas Black motorists would be welcome and outlines the unique dangers of being a Black motorist in the United States. In the thirty years of their publication, Black communities used these books as a harm reduction strategy, describing safer locations to stop and racist treatment by some business and law enforcement officers. These were some of the oldest stories documenting driving disparities. These stories are important early evidence of traffic stop disparities, since modern data collection on traffic stops began only recently, followed by limited attempts to mathematically quantity disparities by interested researchers. While questions of accountability are treated more thoroughly in the Discussion, it is not a stretch to imagine why law enforcement agencies had not prioritized databases of resident interactions, such as traffic stops. Measurement of traffic stops is relatively new because of these accountability dynamics.

In their stead, studies like the work of Epp et al.[39] and Engel and Calnon [38] have used survey data on traffic stops to capture experiences when law enforcement records are entirely missing or insufficient to answer meaningful community and research questions. These studies both report disparities in the experience of people of color when compared to White non-Hispanic drivers, and have the added benefit of narratives and qualitative data that establishes experiences during traffic stops are different, including many kinds of treatment by officers.

Existing research suggestions for measuring racial disparities in traffic stops using stop data share agreement that residential baselines are insufficient, though provide a diversity of solutions to this problem. Many have unmet flaws and limitations. Relatedly, theories for interpreting these police traffic stop rates are insufficiently broad or vague, even if they acknowledge the methodological challenges already mentioned.

One of the earliest attempts at estimating traffic stop disparities was, not coincidentally, completed in North Carolina. Matthew Zingraff et al. completed an evaluation of the North

Carolina State Highway Patrol citation in 1998 [148]. This analysis was only for state highway patrol but contended with similar agency-specific issues since highway patrol areas were broken up into districts. They used a weighting factor, based on estimates of the percent of citations to residents within a district, to improve comparison of residential denominators and tickets collected. Their method was designed to provide drivers as the improved denominator for those districts [116].

Most similar to the aims and methods of this research is work done in Missouri by Jeff Rojek, Richard Rosenfeld and Scott Decker in 2004 [116]. They too acknowledged the issue of residential denominators, and used spatial methods to build a travel-based denominator. However, they did not use race-ethnicity specific data on driving patterns, but chose an inverse distance weighting function and a 20 mile maximum cut off to consider drivers coming into agencies.

Other researchers have advocated very different strategies than attempting to derive a driving denominator. As a first example, Research Triangle Institute (RTI) has created an online tool (RTI STAR), now in use by police departments, based on the work of Grogger & Ridgeway [61]. Acknowledging the challenges in residential denominators and in survey-based approaches to answering those limitations, they recommend a method based on the "veil of darkness" (VOD) approach to assess racial profiling. This VOD method "asserts that police are less likely to know the race of a motorist before making a stop after dark than they are during daylight" [61], and is based on the notion that by constraining only to stops just before and just after sundown the model can describe differences police behavior based on being able to identify from afar the race-ethnicity of a driver. However, this is based on a highly limited notion of potential causes of racial disproportionality – interpersonal prejudice at the time of the potential police stop by individual officers noting the race-ethnic phenotype alone (in this case, skin color) of the driver. Note that this model would fail to identify disproportionality if stop rates were equally high before and after sundown, even if those rates were exorbitantly high compared to white neighborhoods. In contrast, critical anti-racism / white-privilege theory describes racism (and its implied disparities) as structural, where racism and white supremacy operate at these reinforcing, multi-level scopes of influence: (1) internalized in an individual (as racial inferiority and / or superiority), (2) interpersonal interactions and relationships, (3) institutional (e.g. policies, laws, practices), and (4) cultural (norms, symbolism, etc.) (PISB). Instead, RTI STAR limits race disparities to those produced by individual officers making judgements based on skin color (interpersonal), but does not capture, for instance, institutional policies by LEAs that over-police Black and/or low-income neighborhoods day and night, or cultural indifference to these injurious dynamics in the media. It is interpreted by police departments and popular press in a dangerously broad way, e.g.

> *"For the Carrboro Police, the RTI STAR analysis shows a p-value of .8121 for African Americans and .7680 for Hispanics. Both of these values indicate that there is no significant racial bias present. 'For the racial profiling to occur, the p-value would have to be .05 or less," he [Carrboro Police Chief Horton] said. "Ours is much higher as you can see." (Daily Tar Heel, 2018)*

Besides the dangerous overreliance and overinterpretation of p-values [11,60,123], empirical research by Baumgartner [15,15] reinforces this theoretical objection by demonstrating, for example, that even when suppressing data from individual officers with outlier race-ethnic specific stop and search habits (the "bad apple" hypothesis), individual departments often have significant disparities in stop and search habits even when contraband hit rates are similar by those demographics. See the following section on multi-level and structural theories that address these concerns. Ridgeway has also suggested variations on propensity scores as a mechanism for testing post-stop disparities [112], though these methods are difficult to interpret.

Fridell is a prominent author in this space, having published white papers and reports [48] (not in peer reviewed journals) on the subject. She suggests that "researchers should not assume the null hypothesis" on four confounding, potentially causal factors that drive variation in stop rate disparities: (1) residential differences between jurisdictions, (2) differences in driving patterns of those populations, (3) differences in underlying violation rates between race-ethnic groups within and between LEA jurisdictions, and (4) within-jurisdiction differences in driving in "high stop areas." Fridell may imply but leaves implicit variable rates of vehicle ownership and the fundamental truth of cross-jurisdiction driving. When considering racial-ethnic disparities in stop rates, Fridell cautions that "researchers should not assume the null hypothesis," that these factors are the same across race-ethnicity groups. However, Fridell also leaves implicit two additional factors that may differ between driving groups: (5) vehicle access by race-ethnicity and (6) cross-jurisdiction driving. Residential-based rates are at risk of the same errors as her other factors, namely assuming these driving factors are the same between groups of drivers or agencies. Also, Fridell's first four factors do not separate the need for an at-risk, driving-based denominator from disparity rationales. This is in keeping with an implicit definition of disparities that require disparities to be caused wholly and solely by unjust factors, not partly or predominantly. As example, while unsafe driving behaviors may not be the same between groups, there is still a useful basis in generating crash rates based on the same vehicle mile traveled denominators. When separated, studies may choose to treat these potential rationales as confounders and adjust for them. Fridell gives little to no practical guidance on how to better measure these disparities.

Even with accurate stop rates disparities, discussions of their cause and proposed interventions can be difficult. These discussions, as exemplified by the previous example, are often limited to two explanations: either reasonable police response to criminal realities or explicit or implicit[17] personal racial bias by individual law enforcement officers (LEOs). However, as discussed prior, many have noted that racial discrimination operates not only at the personal level, but is also deeply structural, cultural and institutional[17,19]. Agency-specific distribution of police, both spatially and prioritization of certain stop types over others, can drive disparate impacts in the absence of personal bias [16]. Stop type is fundamental to these discussions. Baumgartner et al. [16] included stop categories in their analysis of stop and searches in North Carolina, but few have broken down stop types this way. These stop categories, broadly of three types (moving and safety violations, regulatory and economic violations, and subjective investigation) may be considered interventions that are meant, at face value, to reduce injuries and promote public safety by reducing injuries such as vehicle crashes and interpersonal violence. Variation between law enforcement agencies in injury and stop profiles represent implicit prioritizations of some stop types, and therefore some injuries and crimes, over others.

In “How Police Stops Define Race and Citizenship,” Epp, Maynard-Moody and Haider-Markel [39] agree that experience of people of color differ by traffic stop types, though they focus on sampled survey data to do so. They use this data to suggest that implicit racial bias “are not generally influential but instead are activated by and in the practice of making investigatory stops.” Their research builds on the work of Engel and Calnon [38] that disparities are larger in more discretionary stop types, a finding seen again by Baumgartner et al. in North Carolina.

Conclusion

This study responds to these concerns in the measurement and theory literature directly in the following ways: (1A) using spatial methods to derive more appropriate residential denominators (particularly for sheriff departments, which more typically enforce driving in rural, unpoliced areas of counties); (1B) using data on driving patterns and vehicle availability of those populations; and (3) investigating the relationship of public safety outcomes to variations in traffic stop programs specifying traffic stop types.

2.2 **Relevant Theories, Concepts and Frameworks**

Harm reduction framework

This project attempts to balance both a realistic, harm reduction framework and a sufficiently critical, historical, anti-racist, and visionary lens to policing and public health collaboration. The harm reduction framework acknowledges that this research by itself, and likely no research, can uproot the centuries of white supremacy and racism that have been woven into the histories and present-day legacy of both public health and policing in the United States. Instead, harm reduction advocates addressing symptom severity in a realistic way even in absence of root-level solutions. As example, Black men still make up a stunningly disproportionate number of those incarcerated for drug-related crimes - even though drug use is similar between race-ethnic groups. It is important to acknowledge that “the war on drugs” is a not only a recent and racist legacy of public health and law enforcement collaborations, but also returning to the forefront in as a racist framework option in approaches to the modern opioid epidemic. This research will not stem those tides, and instead is largely trying to ameliorate symptoms of that racist system. It is not directly focused on envisioning the radical, novel forms of community-controlled public safety and enforcement that are viable alternatives to traditional law enforcement structures - essential as those research and visioning activities are.

Anti-racism, critical race theory, and structural determinism

A critical, historical, anti-racist, visionary lens is essential to contextualize this research. Without a critical eye, recommendations from this research may fall into dangerous limitations, including but not limited to: (1) ignoring alternative intervention modalities that don't involve law enforcement and may be less subject to racially disparate impacts and collateral neighborhood and individual harms, and (2) implying additional funding and scope creep for LEAs, expanding law enforcement responsibility to questions of traffic safety, mental health, and public health when alternate, existing, more effective, more specialized, more community-based strategies are underfunded as is. Anti-racist philosophers (e.g. James Baldwin, Franz Fanon, Angela Davis, Cornell West, etc.) and popular education efforts (community dismantling racism organizing [103]) help to establish a broad and nuanced enough model of racism, white supremacy, and policy evaluation. Modern anti-racist policy platforms (e.g. Black Lives Matter's Campaign Zero [24]) advocate for both short-term policy change and deeper alternatives to the policies and practices this project explores.

Public health has also adopted these calls for anti-racism [73,75]. Critical Race Theory [19] and specifically the Public Health Critical Race Praxis [47], call for a structured and critical eye as fundamental to those goals of anti-racism. In keeping with critical race theory principles on the social construction of race-ethnicity, modern literature in epidemiology also calls for a more historical and contextual interpretation of race-ethnicity variables (and experiences by individuals) in models that purport to consider causal inference questions [135].

Related to these principles is acknowledging the multi-level structural determinism of most phenomena, a focus that social epidemiology has explicitly called for (Krieger et al.), in response to a too-narrow focus on individual behaviors as the sole seat of action or intervention. These critical public health frameworks not only acknowledge, but center power differentials as fundamental to understanding health. On power and multi-level theory, Frederick Zimmerman writes, "to ignore power would be to ignore the most important determinant of population health – it would be possible, but it would be theoretically impoverished, ad hoc, and boring" [147]. Interpreting results of these studies correctly and broadly requires these critical lenses. In keeping with a multilevel and structural framework, social epidemiology recognizes a similar distinction in that causes of cases (e.g. "why did this patient get this disease at this time?") are not necessarily the same as causes of incidence in populations [117], and by extension here, the causes of individual stops may not be the same as the causes of the disparities within and between neighborhoods in a jurisdiction. When the unit of analysis is the individual jurisdiction instead of the individual stopped, even the causes of disparities within separate neighborhoods and overall jurisdictions may be different than the causes of variations in disparities between jurisdictions. These may also be considered versions of the individualistic and ecological fallacies, requiring multilevel thinking, grounded in historical and spatiotemporal context [12,124].

Deterrence theory

Deterrence theory in the highway safety framework suggests that increasing traffic control (e.g. lower speed limits), traffic law enforcement, and infraction penalty severity have a reducing effect on traffic crashes [40,113]. However, equally as important are the critics of this theory. The literature on this deterrence impact is mixed, with some studies finding reverse effects or no effect [50]. There are also numerous challenges to its theoretical underpinnings, including that deterrence effects may be significantly non-linear and that generalization of local effects may be suspect. Considering system dynamics lessons, thresholds may exist beyond which the effect of control creates policy resistance and or increased danger. In the unchecked extreme, these policies create an authoritarian police state, with harms greater than those it seeks to prevent.

People over profits

This analysis is performed under the values of socialist critiques of neoliberalism at the macro level [28], most directly the critique of the institutions and cultural dynamics that explicitly or implicitly value profits over people. Instead, I assert that public health should be fundamentally aligned with **people over profits** at both the macro, and in this case, more micro levels. This has concrete implications to the quantitative analysis performed here, where, for instance, when considering the relationship of traffic stop types to traffic crashes and crime incidents that produce injuries, I will weight bodily harm over property harm both theoretically and, where appropriate, quantitatively. For instance, when considering small area models in Fayetteville, crimes and traffic crashes may be combined into a total injury index scored similar to Quality Adjusted Life Years (QALYs), to upweight traffic fatalities, homicide and assault exponentially greater than theft, property damage or suspicious persons incidents. Likewise, when assessing whether crime changed during Fayettevile's intervention period, I will focus on violent crime and UCR index crimes, not all incidents of any time. This framework asserts that while all lives matter (more than property), that must equally include Black lives, and are not allowable collateral damage to activities that seek to protect property - in contrast the precipitating causes of death of many Black men in particular.

Activist scholarship & consequentialist epidemiology

This project aims to follow principles of both activist scholarship and consequentialist epidemiology. Activist scholarship is "the production of knowledge and pedagogical practices through active engagement with, and in the service of, progressive social movements" [125]. By being engaged with ongoing community group efforts seeking to exert influence on and resistance against racist policing models, this research is made more interpretable, receives helpful critique from those who would use it, and finds more opportunities for dissemination and implementation. In this case, this project is privileged to have collaborators in community groups and nonprofits (the NAACP's Orange County Bias-Free Policing Task Force; the Southern Coalition for Social Justice) and select law enforcement agencies aiming to respond meaningfully to community concerns (Fayetteville Police Department). Though there are challenges to being accountable to community groups with real concerns and deadlines and the more abstract, academic knowledge generation process, the benefits are overwhelmingly worth the added difficulty. Along similar lines, consequentialist epidemiology [49] clarifies research priorities, offers perspectives on novel methods, elevates equity considerations, and demands reckoning with the realities of implementation, translation and dissemination. In this instance, consequentialist epidemiology recommends methods robust enough to further the underlying science but direct and developed enough to be able to be implemented with validity by law enforcement and interpreted correctly by the media and public.

In light of these guiding theories and principles, we offer a few specific definitions important for this paper. First, we conceive of racism as (much) more than personal prejudices and stereotypes, more than experiences of interpersonal discrimination whether intentional or unintentional, and more than implicit or explicit personal bias. Instead, according to Camara Phyllis Jones [75],

> *"Racism is a system of structuring opportunity and assigning value based on the social interpretation of how one looks (which is what we call "race"), that unfairly disadvantages some individuals and communities, unfairly advantages other individuals and communities, and saps the strength of the whole society through the waste of human resources. This definition of racism as a system (rather than an individual character flaw, personal moral failing, or psychiatric illness) helps start conversations because we are no longer trying to divide the room into who is racist and who is not."*

The structural focus is particularly important for considering traffic stops, as they are often framed, almost dramatically, as primarily an interaction between one or more officers and a driver, if not also passengers. However, this focus obscures the many ways that structural factors, include history, institutions, and culture, undergird not only these interpersonal interactions but also the fundamental patrol patterns. In a word, per Rose, traffic stops share the reality that the causes of cases are different from the causes of incidence [117]. We focus on measurement and intervention on traffic stop incidence in this dissertation.

CHAPTER 3 - SPECIFIC AIMS

Traffic stop rates are often based on residential populations instead of driving populations and driving patterns that cross multiple jurisdictions. Because of this, though preliminary data suggests already significant racial disparities in traffic stops, these disparities may be widely underestimated. Even with accurate measurement, the link between traffic stop programs and associated public health outcomes is unclear. This project responds to these gaps by the following two main aims: one NC-wide analysis and one focusing on evaluating a traffic-stop-related intervention in Fayetteville, NC. Critical Race Theory [19,46], particularly the Public Health Critical Race Praxis framework, is used throughout the aims and interpretation to drive study design, analysis choices, and discussion of implications.

First, this study aims to establish more accurate traffic stop rates by race for North Carolina police agencies instead of relying on flawed residential population measures (Aim 1A), documenting the degree of difference in measures of disparities as models account for more driving factors (Aim 1B). Driving factors are key to moving from the residential population to the true population at risk of stop in the geographic jurisdiction of a specific police agency. These factors include, but are not limited to, race-specific vehicle ownership, driving frequency and cross-jurisdiction driving patterns like city clustering and distance to work, and are derived from the National Household Travel Survey (NHTS) [133]. These driving factors are combined with the nation's oldest and most complete traffic stop dataset, 20 million North Carolina (NC) traffic stops, including most municipal police departments and rural sheriff departments in NC.

Second, focusing on a single police department in Fayetteville, NC, (Aim 2) a model intervention for reducing racial disparities and reducing traffic injuries by prioritizing moving violation traffic stops is evaluated by comparing against other jurisdiction traffic injury and crime trajectories. Though partly dependent on Aim 1 (for improved estimation of race-ethnic specific stop rates), Aim 2 can partly stand on its own, examining the associated changes in outcome variables relevant for the intervention compared to a control population of similarly sized agencies.

3.1 **Aim 1A: Determine traffic stop rate disparities using driving-based denominators.**

A theoretical gold standard for NC driver driving information would be GPS-linked driving habits and vehicle miles traveled (VMT) of all vehicles in NC and surrounding states by race/ethnicity of driver. Because this gold standard does not, and likely cannot, realistically be obtained, Aim 1A used analysis of the National Household Travel Survey to derive driving-based rates, integrating driving access, driving volume, and trip distance adjustments.

3.2 **Aim 1B: Assess the degree of difference in traffic stop disparity models.**

Because Aim 1 is based on multiple adjustments, and because these adjustment factors are not always as equally difficult to integrate, traffic stop rate, it is instructive to consider the degree of difference between multiple models in assessing disparities. To demonstrate the importance of these iterative adjustments, Aim 1B documented the degree and direction of a demonstrative disparity measure (Black non-Hispanic vs. White non-Hispanic traffic stop rate ratios) will be described through iteratively more complex driving models accounting.

3.3 **Aim 2: Estimate the effects of the Fayetteville Police Department intervention.**

Between 2013 and 2016, Fayetteville Police Department implemented a traffic stop intervention designed to both lessen the racial disparities in traffic stops and simultaneously reduce traffic crash injuries by focusing on safety (i.e. moving violation) stops. Using synthetic control methods to compare Fayetteville's overall intervention impact to that of a control agency constructed from similar agencies that did not enact the intervention, we evaluated the overall intervention impact on disparities and motor vehicle stops, and considered whether there was any change in crime rates due to the reprioritization.

CHAPTER 4 - METHODS

To achieve these two aims, multiple methods were required. Aim 1, improved estimation of traffic stop rate ratios required first deriving race-ethnicity specific driving factors from the National Household Travel Survey [133], then the spatial modeling of VMT distributions by applying those driving factors to statewide residential demographic data from US Census products. Aim 2 used the synthetic control technique to produce weighted combinations of control agencies from a donor pool (large cities in North Carolina), matched on the pre-intervention period for each measure of interest, then compared Fayetteville Police Department to those synthetic control agencies in the post-intervention period.

The study population and datasets and the four main statistical analyses are described below. Three supplemental analyses were used as to explore facets of the main analyses in more detail, though were not included in manuscript chapters. Lastly, this section then finishes with some coding implementation considerations, such as data structures and algorithmic efficiency and speed.

4.1 Study Population & Dataset Details

The populations of interest for Aims 1 are the residential and driving population of North Carolina from 2002 to 2018 and the municipal police and county sheriff law enforcement agencies. The populations of interest for Aim 2 are the driving population of Fayetteville, North Carolina, along with a subset of city law police departments used to evaluate the impact of the

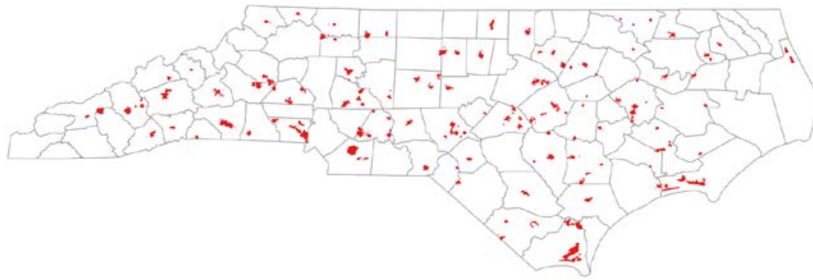
Fayetteville Police Department intervention by deriving a synthetic control department. In order to collect data on or estimate the characteristics of these study populations primary and supplementary datasets are needed.

The primary dataset is the (1) North Carolina Traffic Stop Dataset from 2002 to 2016 from the NC State Bureau of Investigations. Four supplementary datasets are used: (2) the US Census (2000, 2010) and American Communities Survey (2011 to 2016) residential demographic information, (3) the North Carolina subset of the 2017 National Household Travel Survey, (4) Motor vehicle crash data from 2002 to 2016 from the UNC Highway Safety Research Center [131], and (5) Uniform Crime Reporting data from 2002 to 2016 from the NC State Bureau of Investigations [99].

4.1.1 Primary Dataset: NC Traffic Stops

The true theoretical study population is drivers at risk of police traffic stop by a police agency in North Carolina. Establishing a detailed accounting of the study population at risk is Aim 1 of the study.

The primary dataset for analysis is the North Carolina State Bureau of Investigation (SBI)'s database of over 20 million police traffic stops from 2002 to 2016, representing over 300 of the 518 state, county, municipal, campus, and place-specific (e.g. state fairgrounds, capital building) police departments. Preliminary analysis of the agency jurisdiction population coverage suggests this dataset is estimated to include a near-census of over 95% of the all police traffic stops of vehicles in North Carolina in this period.



Data Status	Agency Coverage	% NC pop
Available	9,344,773	98%
Unavailable [red]	190,710	2%

Figure 4.1 Missing data on NC traffic stops is very small

The NC Police Traffic Stop Database includes 308 of 518 police agencies representing all county and most city- and place-based agencies. Though the smallest police departments [in red] are not required to report, the dataset covers nearly 98% of the populated area of NC [in white].

This dataset was created after the state legislature passed Senate Bill 76 requiring data collection for state officers as of January 1st, 2000 and expanded it to include all county sheriff departments and nearly all municipal police departments as of Jan 1, 2002. This includes all state and county sheriff agencies without exception and municipal and place-specific agencies whose jurisdictions either (1) include more than 10,000 individuals or (2) who employ full-time officers at a rate at or above five officers per 1,000 residents . Agency data is available from the NC SBI directly, but is submitted to them at different times by LEAs throughout the state. These LEAs document these activities on a single, consistent state form SBI-122, which is often entered into an agency specific electronic data record. These electronic data systems are agency specific, though there are a few major vendors with a high prevalence of use in NC (the most common being Superion Public Safety Software, formerly Sungard Public Sector). Pilot validation of data with select police agencies suggest that data entry and, in some cases, software errors have created challenges for certain analysis facets (e.g. search reason field has had problems in data transfer to the state), but there are no known or suspected problems in the data elements needed for this analysis: data validity is high with few missing values (e.g. complete race/ethnicity data is missing in 0.1% of the dataset). Our existing relationship with local and state police departments and non-profit users of this dataset allowed us to "ground-truth" these data quality questions. This data from the NC SBI was collected into a more user-friendly website, launched in December 2015 by the Southern Coalition for Social Justice in Durham, NC.

	NC Traffic Stops		NC Resident Population	
	#	%	#	%
Race-Ethnicity				
White *	13,258,385	58.1%	6,223,995	65.3%
Black *	7,076,618	31.0%	2,019,854	21.2%
Hispanic	1,779,330	7.8%	800,120	8.4%
Native American *	181,402	0.8%	108,829	1.1%
Asian *	262,926	1.2%	206,579	2.2%
Other	273,176	1.2%	176,106	1.8%
Total	**22,831,837**	**100.0%**	**9,535,483**	**100.0%**

Table 4.1. Demographics of drivers in NC traffic stop database, 2002-2017

Black non-Hispanic drivers and men are over-represented and White non-Hispanic and Hispanic drivers are both under-represented in the dataset compared to the North Carolina 2010 population. Drivers median age is lower than the North Carolina population by five years.

Preliminary estimates of essential driving variables also available and can be used to crudely adjust these populations to see direction of effect and severity of bias when using a residential population rate. As an example, though the recently released 2017 National Household Travel Survey will be used in this project, the prior National Household Travel Survey (2001) suggests miles driven by race is different between White drivers (mean 12,091) and Black drivers (10,275), and further estimates large vehicle ownership disparities between Whites (84%) and Blacks (51%). Further, given sheriff departments policing rural areas infrequently patrol municipal areas patrolled by municipal police departments, census residential data can be used to provide improved rural populations for sheriff department jurisdictions. Though this crude adjustment method, applying just the vehicle ownership point-estimate as an adjustment across all jurisdiction resident populations yields increased Black-White rate IRRs above 1 in all but 12 agencies, and above 2.0 (twice the stop rate) in 141 of 189 agencies (see Figure).

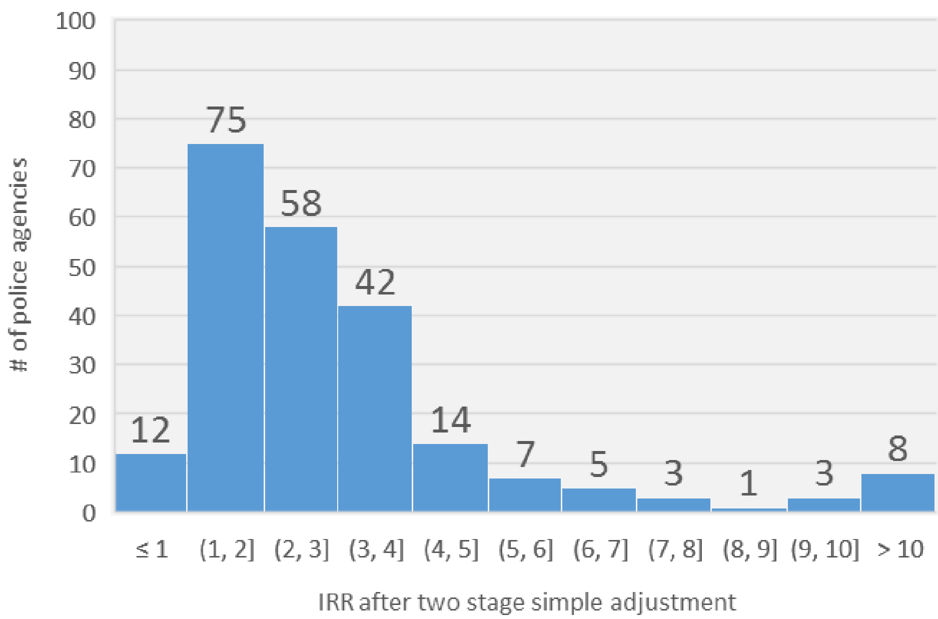


Figure 4.2 Minimal adjustment suggests widespread traffic stop disparities

Only 12 agencies stop Black drivers at or below the rate they stop White drivers, with 141 agencies stopping Black drivers at over twice the rate as White drivers. Includes 189 NC police agencies with more than 1000 recorded stops who had searched both Black and White non-Hispanic drivers. Analysis by Michael Dolan Fliss, UNC-CH, using NC SBI Stop Data, 2002-2013.

This preliminary data, adjusted crudely on only two factors, demonstrate the flaws of the current residential population method. However, vehicle ownership and driving disparities likely vary widely by jurisdiction and should be adjusted differently across space instead of using a single national estimate of vehicle ownership for every NC jurisdiction. Still, nearly all adjustment factors in consideration suggest, by direction alone, that the true rate of stop by driving population and rate of stop by vehicle miles driven are widely disparate by race. Without an integrated model of these adjustments to accounts for spatial proximity, we can only presume that (1) the true stop rates are very different than the residential population-based rates, (2) disparities by race seem to strongly exist, and (3) disparities may significantly increase with confounding control. This level of findings is insufficient for informing community conversations and police policy. To evaluate interpretation of racial bias in policing as they are asked in practice (Aim 2), we need trustworthy, accurate stop rates compare against well-measured community elements like crashes, crime rates and poverty. Preliminary data on the degree of bias in residential-based race disparities in stop rates show a clear need for improved estimates and a system for producing accurate estimates in other jurisdictions.

Therefore, all preliminary data, including the primary police dataset and available estimates from supplementary datasets, suggest that (1) differences in stop rates may be pervasive in North Carolina police agencies and (2) the magnitude of that difference is widely skewed by even single adjustment of known confounders of the resident population / vehicle-miles-driven population relationship. Preliminary data suggested relying on resident populations for assessing race disparities in vehicle stops was significantly problematic, though this measure was widely used in community complaints and agency review and response.

4.1.2 Supplemental Dataset: Census Residential Demographics

Population demographic data for race-ethnicity-specific rate calculations were obtained from the United States American Communities Survey (ACS) and United States census. For Aim 1, US Census 2010 block group data was used to represent the single aggregated estimates for the Aim 1 data period from 2002 to 2018. For Aim 2, year-specific ACS and Census population data for all cities was used, interpolating years 2002 to 2009 using 2000 and 2010 census data when ACS estimates were unavailable. Preliminary analyses were calculated using data downloaded from American Fact Finder, thought the final analysis used the R tidycensus package to programmatically access the US Census API and download block group level demographic data for all census and ACS years. The author requested and was granted a free API key by the US Census to use this service. Demographic data (Race x Hispanic population data) came from ACS table B03002 and Census data from P005.

4.1.3 Supplemental Dataset: National Household Travel Survey (NHTS)

The 2017 NHTS included 8,804 NC households, with information on vehicle miles traveled, vehicles ownership and availability, and race and ethnicity data. The public-use data set can be geolocated more precisely than nearest major city, but the sample sizes stratifying by race-ethnicity would be low for local estimation. The 2017 NHTS was oversampled in North Carolina to produce more accurate state level estimates. NHTS was therefore used to create NC-specific estimates of important driving measures by race-ethnicity, including vehicle access, total VMT by driver, and driving distance distributions. NHTS uses a residential sample frame, by excludes those living in group quarters like military on-base housing, college dorms, and nursing homes/assisted living facilities.

Race-Ethnicity	Measures of Survey Representation		
	Number surveyed	Number represented	Number drivers represented
Asian	307	251,577	184,748
American Indian	156	78,171	57,496
Black	2,444	2,015,261	1,294,804
Hispanic	600	828,660	532,834
Other	522	324,620	199,508
White non-Hispanic	13,556	5,950,650	4,894,298
Total	17,585	9,448,939	7,163,689

Table 4.2. NC survey demographics from the 2017 National Household Travel Survey.

4.1.4 Supplemental Dataset: Crashes

Data on North Carolina motor vehicle crashes since 2002 were obtained from the University of North Carolina Highway Safety Research Center (HSRC) "Create a data table" online web tool, created with funding from the North Carolina Governor's Highway Safety Program [131]. Supplemental analyses that required spatially located points used motor vehicle crash data provided by the Fayetteville Police Department and data acquired through data use agreement with the national Highway Safety Information System (HSIS).

4.1.5 Supplemental Dataset: Crimes

Data on North Carolina index and violent crime data since 2002 were also obtained from the North Carolina SBI public website tool [99]. Both of count and rate data were downloaded, each pre-calculated by SBI. These are standard reporting measures are Federal Bureau of Investigation (FBI) Uniform Crime Reporting data reporting requirements and commonly used a standard measures of crime incidence. Supplemental analyses that required spatially located crime incident points used data provided by the Fayetteville Police Department.

4.2 **Statistical Analyses**

Statistical analysis for both Aims can be broken down into three major components: (1) estimating driving model parameters from the National Household Travel Survey, (2) calculating agency stop rates by applying combinations of those NHTS driving parameters in a spatial simulation to assess the degree and direction of change in a rate ratio disparity measure, and (3) applying synthetic control techniques to assess the impact of an intervention in Fayetteville to reduce disparities and decrease motor vehicle crash fatalities. Each of these analyses is dependent on the previous, since statewide rates are dependent on driving model parameter estimation, and a main measure of interest in Aim 2, the traffic stop rate ratios, are also informed by these driving model parameters.

4.2.1 Driving Model Parameter Estimation (Aim 1)

NHTS data has a complicated nested weighting design, as described in a detailed weighting report [32]. However, the survey comes with prepared weight variables for households (WTHHFIN) and people (WTPERFIN) that are pre-designed to be used for upweighting to produce state-level estimates. This weighting scheme and early results of the driving model parameter analysis was also verified by a phone call with the NHTS survey team in Spring 2019.

Data from the NHTS household, trip, and person tables was joined and upweighted to first produce adjustment factors that would be modeled as single numbers. In order to emphasize the difference between the related concepts of vehicle access, total VMT, drivers, and passengers, separate measures of access and volume were calculated focusing first on the proportion of all residents with driver access, then, given access, the total VMT driven in a year by drivers. All of these measures were derived as averages by race-ethnicity for NC, which required the custom creation of a race-ethnicity variable from the individual race and ethnicity person variables.

	Measures of Access		
Race-Ethnicity	**Household has personal vehicle access (%)**	**Household vehicle use at least a few times a month (%)**	**Any driving during year* (%)**
Asian	99.8	99.0	73.4
American Indian	90.3	95.4	73.6
Black	85.3	88.2	64.2
Hispanic	97.0	97.2	64.3
Other	96.1	97.6	61.5
White non-Hispanic	98.4	98.0	82.2
Total	95.8	96.2	76.8

	Measures of Driver VMT		
Race-Ethnicity	**Annual VMT per driver* (miles)**	**Annual VMT per person (miles)**	**Average miles per trip (miles)**
Asian	8,677	6,372	10.0
American Indian	12,219	8,987	10.8
Black	9,775	6,280	9.7
Hispanic	12,434	7,995	12.4
Other	8,762	5,385	8.6
White non-Hispanic	10,819	8,898	10.4
Total	10,649	8,196	10.4

Table 4.3. NC differences in access and amount of travel (2017 NHTS).

Next, weighted daily trip data was used to calculate an average distribution of VMT at given unidirectional distances, converted to percent of VMT within each radius ring around their residence every single mile up to 400 miles. These raw, exact percentages of VMT within each radius ring were then converted to simple exponential decay linear models, well fit by using the log of the radius multiplied by an interaction term that was a 1 if the radius were under 25 miles, and a zero otherwise. This allowed a inflexion point at 25 miles, and good graphical and statistical fit of these functions. Such a function could then be used for a simple operation for the subsequent spatial model to return, for a VMT catch point at a given distance from a residential point, the percent of VMT to distribute into points at that ring distance. See Chapter 6 for a visual representation of these VMT rings by race-ethnicity.

Confidence intervals were not calculated for either these single number parameters or the modeled VMT distribution functions, though sample sizes as tabulated earlier in the chapter were relatively large for White non-Hispanics (n=13,556) and Black non-Hispanics (n=2,444), the focus of Aim 2.

4.2.2 Residential Attribution (Aim 1)

All models first required consideration of attribution of points to agency patrol areas. Models which used simple adjustment of residential points to prorate into VMT estimates, but did not allow distribution of that VMT at a distance, needed the residential points distributed into agencies patrol areas. Models which allowed VMT distribution needed the VMT catchment grid distributed into agencies patrol areas.

Preliminary exploration of stop patterns, through discussions with police chiefs and sheriffs as well as quantitative stop patterns in the few jurisdictions that maintain point data, suggested that county sheriff departments seldom patrol in municipal areas otherwise policed by municipal sheriff departments. This implies that even residential populations as proxies for patrol areas for county sheriff departments may be meaningfully different from the administrative boundaries.

However, producing appropriate denominators for these rural areas is not as simple as subtracting municipal population totals from police departments, since municipal boundaries may not be entirely contained in any one county. As example, in the case of Orange County, NC (see below), Carrboro, Hillsborough and Efland are entirely contained within the county boundary, but Chapel Hill and Mebane are not. Therefore, small-area census units (e.g. census block groups) were used to re-tabulate jurisdiction-specific residential populations for rural sheriff departments. Again, in Orange County, this not only reduces the population by more than half (from 134,000 to 57,000) in the sheriff control area but also impacts the demographic composition of that rural area of enforcement. In this case, Black residents make up 15% of the rural-only population vs. 12% of the entire county population. Because of the demographic-specific differentials in residential population after this adjustment, traffic stop rate ratios as a means of measuring disparities would change alongside this residential adjustment, even if models that allow for driving distances would reapportion VMT from the cities into counties and vice versa.

	Census Population		Rural-Only Population	
	#	%	#	%
Total	133,801		56,986	
White nH	99,495	74%	43,963	77%
Black	15,928	12%	8,429	15%
Hispanic	11,017	8%	4,673	8%

Source: 2010 US Census, Orange County, NC

Figure 4.3. City and county jurisdictions and population adjustment near Orange County, NC.

City police departments with few county sheriff patrols emphasis may be (left) entirely encapsulated within the county (e.g. Hillsborough, Carrboro) or may cross county boundaries (e.g. Chapel Hill, Mebane). The rural portion of the sheriff jurisdiction's population may differ in both number and demographic distribution when compared to the overall county distributions (right).

4.2.3 Spatial Simulation (Aim 1)

After driving adjustment factors and functions were calculated and residential and VMT catchment grid points were distributed into city and county agencies, these parameters and spatial objects were used in a spatial simulation to derive and compare traffic stop rate ratios (TSRRs) under multiple model assumptions.

To address racial-ethnic differences in access to vehicles, volume of annual driving, and driving through multiple agency patrol areas, we use NC-specific statewide estimates from the 2017 National Household Travel Survey (NHTS). For access, 82% of White non-Hispanic people and 64% of Black non-Hispanic and Hispanic people have access to a vehicle as a driver in NC. For amount of annual driving, NHTS suggests 10,819 VMT per year for White non-Hispanic drivers, 9,775 for Black non-Hispanic drivers, and 12,434 for Hispanic drivers. These single value adjustment factors, alongside others of interest, are included in Table 2. To model travel between agencies, we use NHTS vehicle trip data to find, for example, average trip distances for White non-Hispanic drivers was 10.4 miles, Black drivers 9.7 miles, and Hispanic drivers 12.4 miles. The more detailed distance decay functions were used instead of these single-factor average trip distance estimates.

For quantification of these dynamics, these access, amount, and multi-agency distribution estimates are transformed into parameters used to support spatial models of race-ethnicity-specific VMT distribution, traffic stop rates, and subsequent traffic stop rate ratios. Nine models are evaluated, each adjusting zero, one, two, or all three driving factors: zero adjustment models include (1) a residential count model representing the status quo practice and (2) a driving transformed model where all residents travel the same 10,000 VMT a year; single adjustment models include (3) multi-agency driving adjustment only, (4) adjustment to amount of driving only, and (5) adjustment to vehicle access only; (6-8) double adjustment models include all pair-wise combinations of models 3, 4, and 5; and (9) a single model with all three adjustments. In all models, driving-points are uniquely assigned to patrol areas as described previously in keeping with common patrol overlap realities (e.g. sheriffs are not assumed to patrol their entire counties equally if cities are patrolled by municipal police departments). While sheriff departments may use the entire county for rate calculations, study interviews with police chiefs and sheriffs and limited supplementary GPS data suggest this adjustment is closer to the realities of patrolling.

Residents were modeled by US census 2010 counts of people attributed to census block groups, the second lowest level of spatial granularity. These residents are then prorated by access parameters into drivers, transformed by driving volume parameters into VMT estimates, then distributed over space using a unidirectional spatial density fall-off function based on the proportion of trips within each distance ring.

VMT was distributed into a 1-mile square VMT catchment raster grid (53,818 points uniformly distributed across the state, subsetted to the points within the state boundary) based on distance from each block group centroid to the raster point. Each point is assigned the best-match patrol area: city police departments patrol within municipal boundaries, and sheriff departments patrol county areas not patrolled by police departments. State highway patrol was not modeled in this analysis (see Discussion).

After VMT totals for each model are attributed to catchment grid points, and those point totals are aggregated into agency VMT totals, models are then standardized against a single DOT-estimated VMT total by proportionally transforming each so the total VMT for the entire system is the same 1.1 billion VMT per year regardless of model (Perdue, 2010). This standardization not only ensures model TSRR estimates are comparable but is reasonable given only one consistent VMT total was experienced by the system.

The agency-specific stop rate estimates, after modeling their rate denominators in multiple ways, are then treated as the unit of analysis to consider the direction and degree of change in the race-ethnicity-specific difference for city police and county sheriff law enforcement agencies. The distribution of agency TSRRs are combined without weighting, e.g. the distribution of IRRs is described regardless of agency jurisdiction size or number of stops of the agency. While all city and county law enforcement agency estimates were modeled, agency estimate distributions were filtered to only include 177 agencies with patrol residential populations great than 10,000, complete data over the study period (2002-2017), and at least 1,000 stops over the study period.

Power and sample size considerations

Determining agency-specific stop rates through simulation does not follow a conventional model of hypothesis testing or effect estimation, and so traditional power considerations, testing, and estimation do not apply in the same ways. Power estimation is most appropriate when sampling or small number concerns limit the magnitude of the detectible effect. In this case, for Aim 1, we have a database with near census-level data coverage of almost 20 million police stops and are aiming to accurately estimate agency-specific stop rates, not effect estimates.

However, estimating the total and race-specific stop rates through the proposed simulation does create precision implications that need to be addressed. Current stop rates based on resident populations are biased and may limit model results in at least two ways. The first bias comes from missing data due to (A) smallest police agencies not being required to report and (B) stop data not including checkpoint data statewide. We believe this bias is the smallest, since missing data from smaller police agencies will not bias agency specific stop rates and should not significantly bias statewide estimates. Anecdotal evidence suggests not recording checkpoint stops may bias estimates of race disparities toward the null, suggesting disparities may be greater if they were included. They are believed to make up a small percentage of total stops.

The second bias, seemingly the largest one and the subject of Aim 1, is from ignoring race-specific driving data and not using appropriate populations at risk and vehicle miles traveled. Agency specific adjustments for these driving variables are not available, so must be modeled through multiple simulation, which is an opportunity to close this bias at the cost of precision. These driving covariates will be modeled deterministically in this analysis, though could alternately be modeled probabilistically by a distribution around point estimates from other data sources. Assessing power from this model requires building the entire model to enable spatial interaction and running on simulated data, so is not possible a priori. However, given that crude adjustment shows order of magnitude changes in stop rates, this closure of bias from residential-based stop rates should far outweigh the loss of precision of by estimating using a modeling approach, the only feasible method of reducing this bias given the lack of gold-standard communing measurement by race-ethnicity and jurisdiction.

Agencies with the smallest stop rates, though near-complete registries and not samples, may still be inappropriate to produce agency-specific estimates of racial disparities. However, because the dataset already exempts the smallest agencies from submission, these smallest-stop-rate agencies are rare in the dataset, allowing agency-specific estimates of 177 of the 308 county and city agencies. This will be sufficient for modeling in Aim 1A and meet the agency-specific racial disparity measures goals in Aim 1B.

4.2.4 Synthetic Control (Aim 2)

While simple Difference in Difference (DiD) modeling can compare the before and after trajectories of policy outcomes to consider the total difference of a policy effect after a time point – in this case, the effect on traffic stop demographics and traffic injury rates. As example, below are FPD's annual safety-related percentage of all traffic stops from 2002 to 2015, demonstrating the real (i.e. with intervention) and counterfactual / theoretical status quo trajectories. Control agencies (e.g. Raleigh, PD, in below graphic) that did not enact similar interventions can also have their outcomes modeled as if they had enacted these interventions.

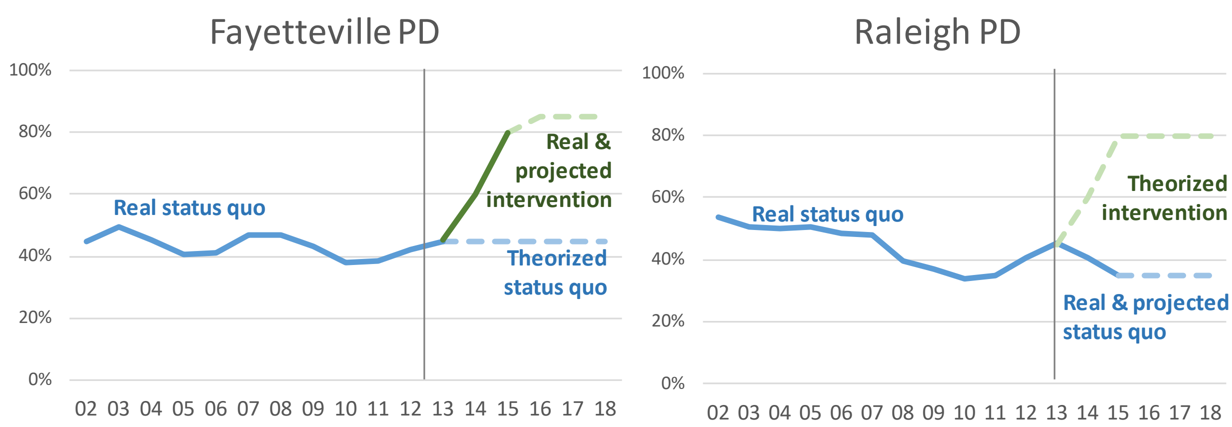


Figure 4.4. Safety-related percent of traffic stops in DiD model, Fayetteville and Raleigh..

However, DiD modeling has limitations, including the often violated assumption of parallel trends between the pre-intervention period and the post-intervention counter-factual. As an example in this case, DiD modeling would require the assumption that the trend for many driving parameters before 2012 would have remained the same during the intervention period, between 2012 and 2016. However, state trends such as continued post-economic recession recovery may have changed the amount of driving (and so traffic stop or crash trends) in all agencies, including Fayetteville. Without better controls, estimates of Fayetteville's post-intervention change would wrongly ascribe statewide trends to Fayetteville alone. And single agency controls may not be similar by themselves to provide adequate match in the pre-intervention period and comparison in the post-intervention period. To answer these limitations, synthetic control techniques been developed, recently joining social epidemiology tools as a method of estimating policy intervention effects in ecological, observational data. Synthetic control techniques model a theoretical control unit using match-weighted data from other control units and has been specifically used before in injury epidemiology [81].

Matching methods in synthetic control techniques vary in simplicity and capacity: pre-intervention outcomes can be matched one at a time or concurrently and can be simultaneously matched on zero or more covariates. For Aim 2, time-varying pre-intervention data was only modeled and matched on each outcome of interest. The improves on matching against time-invariant outcomes or covariates alone, e.g. the pre-intervention average or sum of time-varying observations on one or more pre-intervention variables (e.g. the average percent of safety stops, or the sum of all fatalities). This vector of agency-specific weights for each measure of interest was determined such that the pre-intervention match covariate in the intervention agency are exactly if possible or approximately if necessary close to the same covariates weighted by a linear combination of control agencies from the donor pool, each with a weight greater than or equal to zero and sum of those weights equal to 1.

In practice with small numbers, this technique finds 1 or more agencies that, in linear weighted combination, generate synthetic agencies with a pre-intervention trend that maximizes similarity against the intervention agency (or units, in larger studies) on a specific outcome measure. These weights, determined by the pre-intervention period, are then applied in linear combination to the post-intervention period, and differences compared in the intervention and synthetic control agencies compared to generate an estimator of the comparison between the intervention and the counterfactual intervention block as if it did not receive the intervention. When using multiple covariates (not done in this analysis), researchers may allow the algorithm to decide the weight of the covariates for matching if content knowledge allows or allow this to be done programmatically. Abadie and Gardeazabal [5] provide a simplified matrix math approach to these linear weight combinations in (2003) and a review of the technique with examples in an statistical article [3] accompanying the Synth R package. The final table of weights from each individually modeled synthetic control are below. These weights are carried over to the post-intervention period to compare to the intervention agency's actual values. See Chapter 8 results for the derived table of agency weight vectors for each measure.

The microsynth R package provides three supplemental methods for statistical inference, estimation of variance, and associated confidence intervals: Taylor series linearization, jackknife, and permutation methods [114]. In each case, the point estimate and associated confidence intervals are separately estimated from the synthetic control modeled post-intervention annual average and annual percent change. We chose Taylor series linearization for estimates of confidence intervals because of the relatively few units that would limit resampling- and permutation-based methods. Given the number of units, these point estimates may not exactly match those derived from the synthetic control weighting-based method and therefore may be unsymmetrical: we report both the percent change from synthetic control modeling and the Taylor series linearization method approximation of the same to assess 95% confidence intervals.

4.3 **Supplemental Analyses**

Three supplemental analyses were undertaken to answer substantive questions about Aims 1 and 2, though left out of manuscript chapters 7 and 8. First, in order to both provide preliminary driving model parameters before NHTS analysis was complete and to explore the robustness of these simplified, unidirectional VMT models, the author captured all exact drips of his vehicle during 2017. Second, a simplified deterministic tree model was used to explore the overall effect of the Fayetteville intervention using simpler statistical techniques. Lastly, to better understand the implementation of the Fayetteville intervention, multiple analyses were completed at the sub-agency level using Fayetteville's (rarely available) GPS data on traffic stops and corresponding (commonly available) GPS data on vehicle crashes during this period of time. This last analysis was extensive and may warrant future papers; therefore, the bulk of this analysis is included as an Appendix and only summaries are provided in this chapter.

4.3.1 Gold Standard, n=1: comparing unidirectional NHTS model to author's driving

As a practical demonstration of the extent of this limitation, using an inexpensive on-board diagnostic (ODB) tool that plugs into a car engine's computer, the author tracked all exact driving paths (1,336 trips) he took in his single vehicle over the course of an entire year, then downloaded and processed that spatial data in R (below, visualized against NC counties). Though centered in his residence of Chapel Hill, NC (with driving enforced by the Chapel Hill Police Department), because of the realities of work and community activity space, he regularly contributes significant portions of his annual VMT to jurisdictions patrolled by nearby municipal police departments of Carrboro, Hillsborough, Durham, and Raleigh cities and rural areas patrolled by Orange, Durham, and Wake County Sheriff Departments. In contrast to both his residential location (Chapel Hill, in Orange County) and the bulk of his VMT at-risk time, his only traffic stop was for a speeding violation in Clemson, South Carolina, a place he visited twice that year where he contributed less than 1% of his VMT.

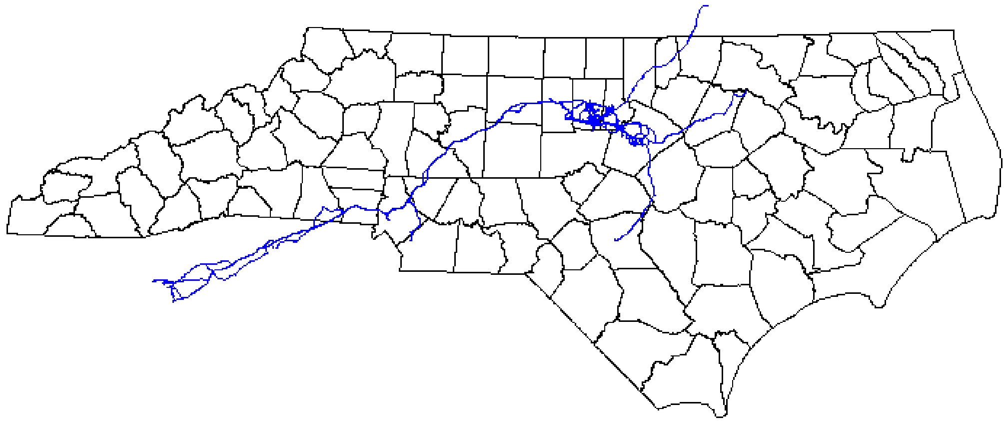


Figure 4.5. A census of the author's driving trips, 2017 (N=1,336).

Most of these trips were centered around his residence in Chapel Hill, NC and work activity spaces, but also includes VMT in over 20 county sheriff department and dozens of municipal police department jurisdictions in NC. He contributed different quantities of total vehicle miles traveled by race-ethnicity (White non-Hispanic) status to each of these jurisdiction's stop rate denominators.

A unidirectional model was built based on this driving pattern by collecting total VMT and the VMT distributed at different radius rings from his residential location. In contrast to 2017 NHTS data where White non-Hispanics drove on average 10,819 miles, the author drove 9,568 miles. These trips were also shorter on average (see figure), so had a less skewed unidirectional radius distribution.

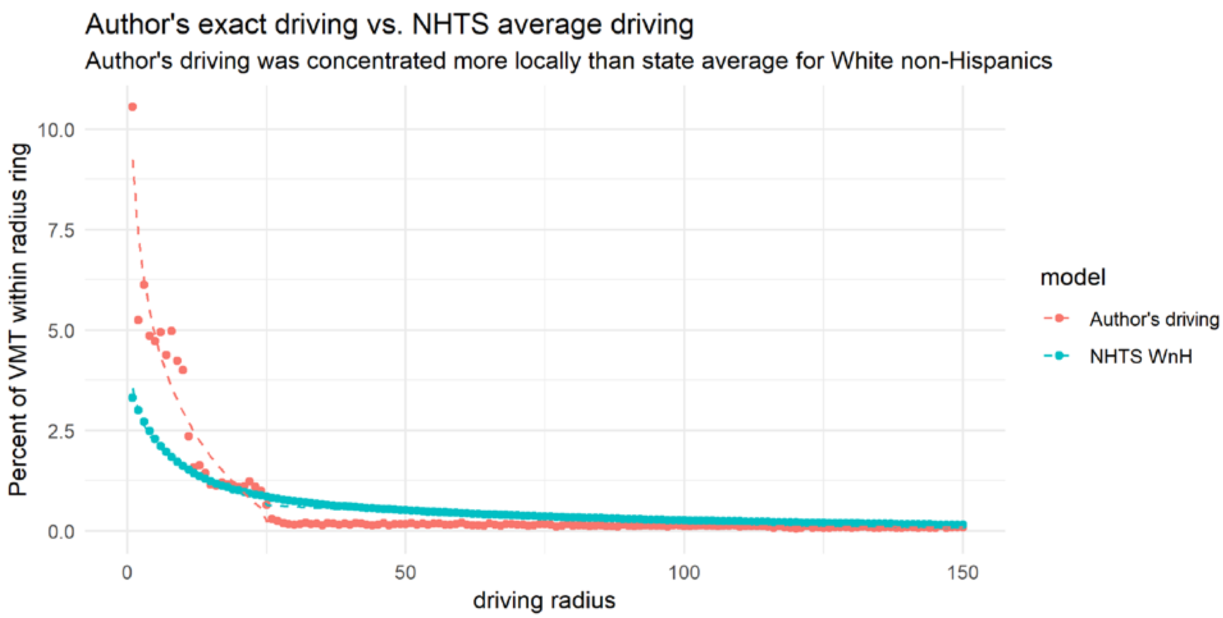


Figure 4.6. Author's VMT distribution function vs. NHTS White non-Hispanic drivers.

Focusing on counties for graphical simplicity's sake (not agencies, the eventual unit of analysis), and though the author's driving is not representative of all (any other) drivers, results suggest the basic residential model that distributes 100% of his VMT to Orange County accurately apportions only 70% of his VMT to the correct county (see below).

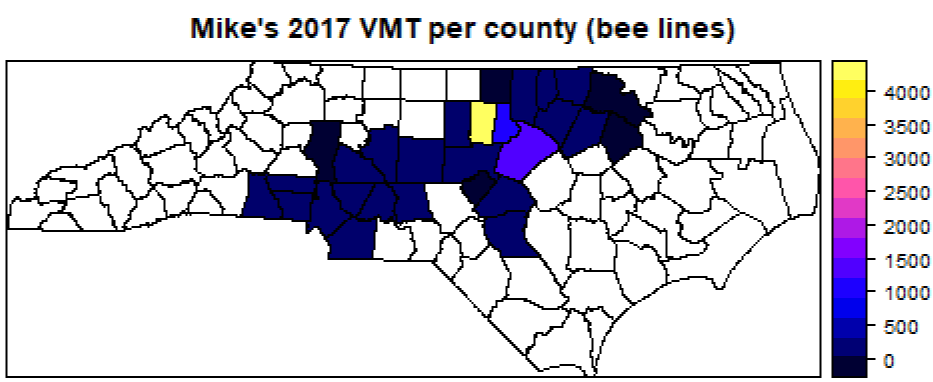


Figure 4.7 Author's VMT correctly distributed into 28 North Carolina counties.

These exploratory analyses suggest a unidirectional model, while imperfect giving clearly directional travel, can still be relatively accurate, however. This is because the largest amount of VMT is distributed into the closest agencies: in this case, not only Chapel Hill (and Chapel Hill Police Department), but the Orange County Sheriff patrol area and neighboring counties. This model would still distribute VMT to counties and agencies the author never visited. It would also miss the dynamic that the author's commute to work in Raleigh and Wake County mean he's contributing more VMT there than Durham. This dynamic is because commuting to work through Durham efficiently contributes a minimum of VMT on major highways (e.g. almost entirely patrolled by State Highway Patrol, not Durham Police Department or Sheriff), whereas work end-points in Raleigh also experience drive around for lunch and meetings. Also, importantly, while this may capture VMT with some accuracy, it notably misses distributing any VMT into the (SC) agency where the author received his sole ticket. This can be thought of as a kind of misclassification of exposure and outcome space, in that analyses for South Carolina would include my ticket, but not my VMT. However, these errors are bi-directional, so some other driver may likewise have contributed a proportionally small amount of VMT to Chapel Hill, and received a ticket. Rate calculations only require accurate assignment of numerators and denominators, and race-ethnicity-specific VMT denominators do not have to be contributed by the same driver that received the ticket, only that the quantities are correct.

These considerations suggest that while unidirectional models have numerous specific limitations, they likely still represent an improvement over effectively disallowing any driving between or across jurisdictions.

4.3.2 Exploratory analysis and alternate methods for Fayetteville intervention effects

Decision tree analytic models are a system analysis method used in health policy and health economics to apportion outcomes, such as monetary costs or quality-adjusted life years (Gold et al., 2009). Similar to conditional probability trees, decision nodes are populated with evidence-based probabilities, that, given the conditional location on the tree, describe the likelihood of moving to a given branch. These tree-based methods can also accommodate non-deterministic, probabilistic risk distribution intervals.

Though simplified for this purpose, these tools can be appropriated to describe policy pathways. Below is a decision-tree model describing the status quo and counterfactual intervention pathway in stop types and driver demographics for Raleigh, NC, as a means of operationalizing the above right graph of its possible and theoretical pathways. This model (M1 in following figures), presents an alternative where the demographics of those stopped for a given stop type remain the same, but the proportion of those stops changes (to 80% safety, similar to Fayetteville's eventual distribution) under the reprioritizing intervention.

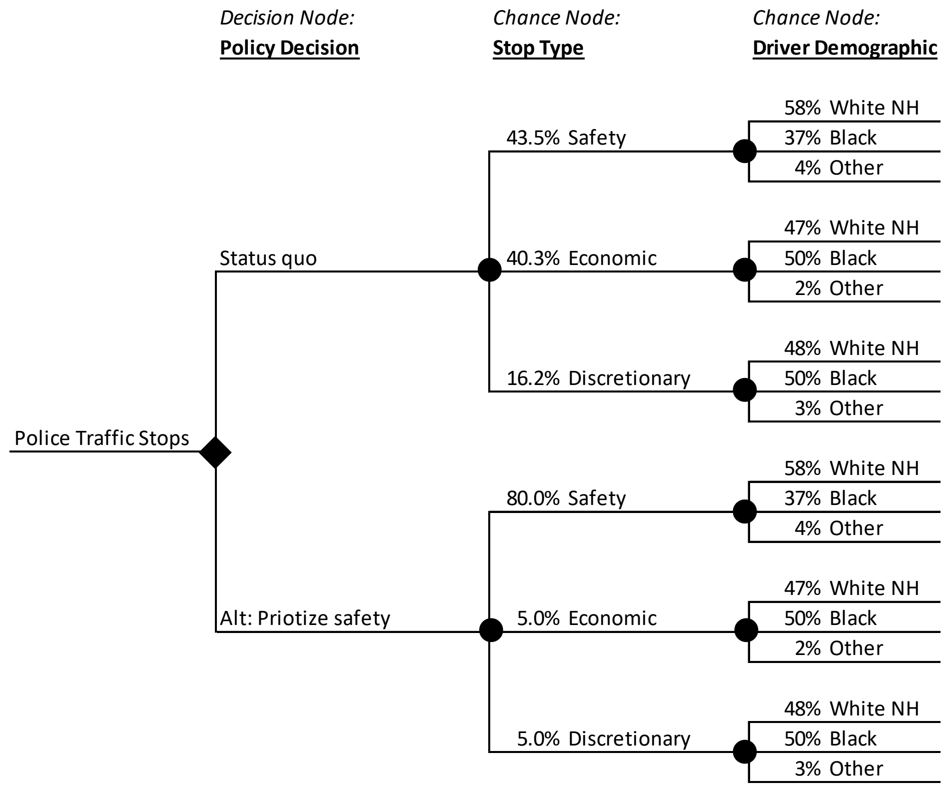


Figure 4.8 Decision-tree-like structure of policy decision evaluation.

Example tree from Raleigh, NC, 2015, though each agency was modeled using their own prior data. This represents Model A1, using existing safety-stop demographic balance instead of adjusting for any alternatives.

Outcomes from a simplified deterministic simulation using these policy-specific decision tree models, with simple accounting for different demographic distributions per stop type given different demographic balance hypotheses, are in the below tables. These suggest that Fayetteville may have had some reduction in the number of drivers stopped that are Black, and that the top fifty largest PD and county sheriff departments may have similar outcomes.

Though economic and subjective stop reasons are tied to low-income neighborhoods that are disproportionately Black and Hispanic, there is little evidence that drivers drive commit moving violations at significantly different rates. Further, residential population-based metrics may under-report disparities since vehicle access is different by income and racial-ethnic identity3,4: e.g. 84% of White and 51% of Black households have access to a vehicle. This suggests that safety stops may still be disproportionate to an underlying residential or driving based benchmark.

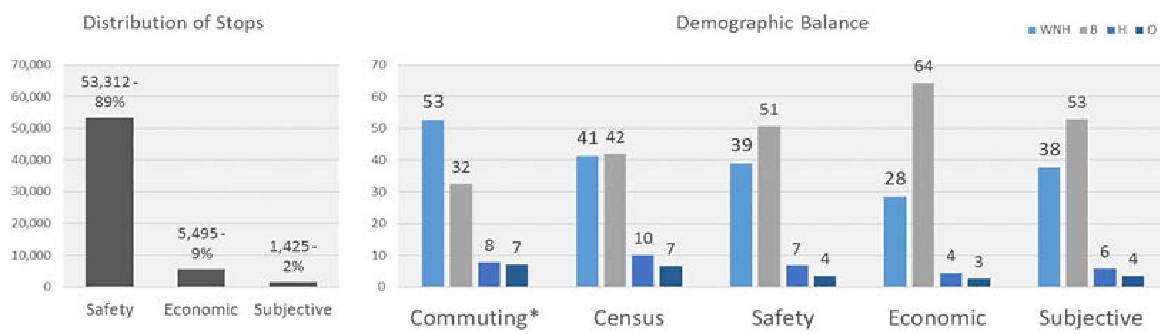


Figure 4.9 Fayetteville distribution of police stops, demographics, and three stop types..

Therefore, two additional policy models are explored, each a ramp up or down to a 2015 racial distribution of safety stops based on the below benchmarks: Model A2, where the safety stop percent mirrors the residential demographics of LE agency jurisdiction instead of as observed, and model A3: % Safety stops mirrors simplified driving demographics of LE agency jurisdiction.

Input Parameter	Modeling Basis	Source
# of stops	Actual & Alternatives	Historical: 3 previous years' data, LE specific
% of stop types	Actual	Previous year's data, LE specific, carried forward
	All alternatives	2013, 2014 trend to 2015 target: 80% / 15% / 5%
% by race by stop type	Actual	status quo: previous year
	A1: Fayetteville Policy	Previous year's race-specific stop %
	A2: Census Target	2010 US Census
	A3: Commuting Target	National Highway Transportation Survey

Table 4.4 Input parameters for decision tree models.

Input parameters are varied based on law enforcement agency and year-specific stop profile, stop-by-race percentages, and, in one-way sensitivity analysis models A2 and A3, their specific residential and driving demographic profiles.

Fayetteville's change in policing strategy in 2013 (A1) is anticipated to reduce the percent of stopped drivers who are black from 57% to 54% over five years, representing 5,700 fewer Black driver stops. If the 50 largest agencies followed Fayetteville's lead, they would reduce the percent of black drivers stopped by 3%, representing nearly 100,000 fewer stops of Black drivers and less than a 1% change in Hispanic drivers.

Though preliminary, this suggested that though reprioritizing safety-related stops was a promising strategy for reducing racial disparities in law enforcement traffic stops, existing disparities in safety-related stops limit the effectiveness of this intervention on reducing the disparate impact of policing in communities of color. Though reprioritizing safety-related stops is a promising strategy for reducing racial disparities in law enforcement traffic stops, existing disparities in safety-related stops limit the effectiveness of this intervention on reducing the disparate impact of policing in communities of color. Institutional-level policy change, not just through individual racial bias training of police officers, may reduce racial disparities in stops.

LEAs interested in further reducing racial disparities in traffic stops may be able to reduce them by (1) increasing the proportion of safety-related stops, (2) critically examining racial disparities by stop type, (3) using appropriate estimates and benchmarks of populations truly at risk of stop, not just residential populations. Though I did not explore estimates of the impact on traffic crashes and injuries using these methods, they too could follow similar analysis methods beyond the primary analysis using synthetic control.

Fayetteville PD Stop Purpose Policy Scenarios, 2013-2017

	% Drivers of Total Stops that are			# of Drivers Stopped that are		
	W N-H	B	H	W N-H	Black	Hispanic
Status Quo (3)	34%	57%	6%	71,462	119,155	12,410
Alternative Scenarios						
A1: Fayetteville Repriotiziation	36%	54%	7%	75,143	113,487	13,622
Diference vs. Status Quo Policy	**+1.8%**	**-2.7%**	**+0.6%**	**+3,681**	**-5,669**	**+1,212**
A2: A1 w/ demo balance (*)	38%	48%	8%	79,659	101,323	17,595
Diference vs. Status Quo Policy	**+3.9%**	**-8.5%**	**+2.5%**	**+8,197**	**-17,832**	**+5,184**
A3: A2 w/ commuting adjustment	45%	43%	7%	94,071	89,284	14,693
Diference vs. Status Quo Policy	**+10.8%**	**-14.3%**	**+1.1%**	**+22,609**	**-29,871**	**+2,282**

A1 represents the Fayetteville (real) policy choice and its projection to 2017 of going, over the course of three years (2013-2015) to 80% safety stops, 15% economic stops and 5% subjective stops (from 42%, 46%, 11% respectively). A2 represents a (theoretical) shift in the demographics of safety stops only to Fayetteville's residential demographics from 2013 to 2015, given Fayetteville's population of 42% White non-Hispanic, 42% Black, and 10% Hispanic. Model A3 represents a (theoretical) shift in safety stop to mirrors Fayetteville's estimated driving population, accounting for large differences in household vehicle access, estimating Fayetteville drivers as 53% White non-Hispanic, 33% Black, and 8% Hispanic.

50 Largest City PD and County Sheriff Stop Purpose Policy Scenarios, 2013-2017

	% Drivers of Total Stops that are			# of Drivers Stopped that are		
	W N-H	B	H	W N-H	Black	Hispanic
Status Quo (3)	46%	43%	8%	1,440,799	1,368,222	247,990
Alternative Scenarios						
A1: Fayetteville Repriotiziation	48%	40%	8%	1,523,146	1,272,666	246,097
Diference vs. Status Quo Policy	**+2.6%**	**-3.0%**	**-0.1%**	**+82,347**	**-95,556**	**-1,893**
A2: A1 w/ demo balance (*)	50%	36%	9%	1,575,743	1,136,858	296,686
Diference vs. Status Quo Policy	**+4.3%**	**-7.3%**	**+1.5%**	**+134,943**	**-231,365**	**+48,696**
A3: A2 w/ commuting adjustment	56%	32%	8%	1,753,083	1,007,223	248,707
Diference vs. Status Quo Policy	**+9.9%**	**-11.4%**	**+0.0%**	**+312,283**	**-360,999**	**+717**

Table 4.5 Traffic stop demographic changes under a deterministic, agency-specific model.

Fayetteville did not maintain status quo policy, but instead did enact the intervention, and the 50 largest cities did maintain the status quo, so their intervention dynamics are modeled based on historical distributions and projected to 2017 (analysis completed in 2015).

4.3.3 Small area examination of Fayetteville intervention implementation

Regardless of Fayetteville's overall effect (Aim 2), the small area effects may vary; assuming they are the same is a geospatial version of the atomistic / ecological fallacy (Subramanian et al., 2008). A sub-jurisdiction analysis is also a means to test the validity of the intervention itself, and answer questions about the implementation.

As example, FPD officers may or may not have actually followed the new policy directives in a way that establishes a reasonable mechanism of effect: they may have increased safety traffic stops specifically in high crash areas or instead may have clustered safety stops in areas seemingly less useful to crash prevention. As a second example, while at face value the intervention reduced the percent of stops of Black drivers, that demographic effect may be different over space within that jurisdiction.

We used two complementary methods of spatial analysis to understand these dynamics: container-based, area-level analysis that accounts for neighborhoods, and a spatial field / surface model robust against arbitrary administrative boundaries. The details of these analyses are beyond the scope of this main dissertation and relegated to Appendix 3. However, the exploratory analysis suggestive some important context for the Aim 2 results.

First, the intervention definition is more complex than what is summarized in Aim 2. Considering the area-level trends, the percent of safety stops did dramatically increase, signifying the implementation of some change process. However, the raw number of total stops first sharply dove, then increased. Discussion with law enforcement administrators suggested this directive was to de-prioritize non-safety stops, prioritize safety stops, and use data to cluster those stops around areas that had high crashes. These first two directive represent a culture shift for Fayetteville, and again, per administrators, officer response was mixed. Some officers elected to leave, others may have been let go, and others needed retraining. Some may have chosen not to document stops in general or specific types of stops, and some may have chosen not to reduce their stops. Broadly, though discussed elsewhere, this is the first component of the Ferguson effect: when officers respond to perceived community mistrust and demands for accountability by reducing their output.

However, the sub-spatial analysis in Appendix 3 suggests the intervention period did correspond with new or changed activities happening at the sub-agency level. Multiple analysis types suggest that safety stops in particular may have clustered moreso around higher crash areas. This provides evidence that while there may have been some resistance to the intervention design, over time there were activities that, through a reasonable mechanism, may have corresponded to the changes seen in Aim 2 results.

Sub-agency GPS data then provides multiple benefits in this case. First, the use of GPS data of both crashes and stops may have uniquely allowed administrators to direct the intervention as it occurred (e.g. crashes are here, send more patrol there), leading to better outcomes. Second, it allowed administrators to evaluate whether their intervention plan was being followed, giving them the tools to understand the operation of their employed officers and patrol patterns. Third, it would allow post-intervention analyses (such ones explored in Appendix 3) to summarize their sub-agency effects; future manuscripts may document these changes in stops and crashes at the sub-agency level for Fayetteville. Without GPS data, law enforcement administrators may not be able to as effectively direct “hotspot” type programs or evaluate their effectiveness.

4.4 **Coding techniques, data structures, packages, and algorithm efficiency**

The simulation in Aim 1, while relatively simple to explain, is complex to implement efficiently. It requires multiple model parameters operating in multiple models, and these model parameters include not only constants, but VMT distribution functions and underlying spatial data. Aim 2 is less complex, but still involves multiple models and would benefit from efficient implementation. This section details these data structure and algorithm considerations

First, a geospatial polygon surface representing LEA patrol areas must be created to act as a catchment for these vehicle miles traveled. For aim 1, this polygon tessellated surface includes county sheriff departments in rural areas, municipal police departments based on city boundaries. State highway patrol, though they could be based on state road buffers, operate differently than city and county agencies and are left out of this analysis. Hospitals and universities are likewise currently left out. This surface can be created by aggregating the centroids of a smaller-unit surface, effectively spatially punching cities through county patrol areas by aggregating small-units first into cities, then only those unassigned to municipal agencies into counties. For efficient testing purposes tracts are used, though the final results benefit from more precision by using block groups. Centroids of this small areal unit then each have their by-race-ethnicity population data used, alongside between zero and three travel adjustments from NHTS. Total access and volume adjustments are derived at the state level and treated as a constant for population data. The algorithm efficacy [84] of this step is therefore O(N), where N is the number of small-area residential units used to tesselate the plane (e.g. ~2,000 census tracts, or ~6,000 census block groups). The final adjustment of distributing VMT over space requires race-ethnicity specific VMT distribution functions derived from NHTS. These distribution functions take a single distance parameter and return the percent of VMT at that distance ring. Each residential center point is then used as an origin to distribute VMT into a VMT catchment grid, set as a 1-mile square VMT catchment raster grid (53,818 points) uniformly distributed across the state and clipped to North Carolina boundary. The algorithm complexity of this VMT aggregation step is effectively O(N x M), where N is again the number of small-area residential units used to tesselate the plane and M are the number of VMT catchment points.

R was used for all statistical analyses, which offers many unique benefits. The tidyverse [108] package for data manipulation and ggplot2 [141] plotting provides tight recoding and the and implementation of the grammar of graphics [140,144] for consistent visualization framework. The sf package [105] implements the simple features format for handling spatial objects, allowing integration of spatial and non-spatial analysis techniques in the tidyverse framework. The microsynth [114] package was chosen as it implemented the newest synthetic control model techniques including the ability to generate confidence intervals. Notably, the Microsynth package was designed for synthetic control studies with large numbers of units (unlike the relatively small number of agencies in Aim 2), and if it weren't for the desire to generate confidence intervals, the older Synth package may have sufficed. The rvest [142] enabled data scraping from websites. The tidycensus [137] package offers programmatic access to the US Census Application Programming Interface (API). The purrr [67] and furrr [136] (future purrr) packages allow R to perform distributed computing, such as over multiple processors or multiple machines.

Without changing this overall algorithm, the analysis takes many hours to complete on a high end, personal laptop. However, while small algorithm efficiencies can improve this order O(N x M) performance in practice. First, the distribution of VMT beyond 100 miles is negligible, so the distance radius cut off for VMT distribution can be modified for testing (e.g. 15 miles), efficient estimation of final results (e.g. 50 miles), or complete analysis (no radius limits). Second, R can use its more advanced tibble data structures (tables that allow list-columns) to retain complex model data (e.g. rows of model parameters, functions, spatial objects, etc.), shrinking floating global variables and increasing coding efficiency. Lastly, O(N x M) is model specific, and Aim 1 iterates over 9 model combinations. For nearly a 10-fold increase in algorithm speed without reducing complexity, functional (e.g. non-iterative) programming techniques can be used from the purrr [67] package alongside the map-reduce framework from the furrr [136] (future purrr) package instead of an iterative loop. This allows the analysis load to be distributed over, for instance, all machine available on a cluster, or in this case, all processors on a local machine. Aim 2 is less algorithmically complex, but still benefitted from functional programming techniques from the tidyverse, such as the flexible tibble-based data structures to store models and results efficiently. These techniques enabled faster testing and development of the analysis.

CHAPTER 5 - A PUBLIC HEALTH CRITICAL RACE PRAXIS VIEW OF LAW ENFORCEMENT TRAFFIC STOPS

5.1 **Overview**

Conventional frameworks suggest traffic stops promote public safety by reducing dangerous driving practices and non-vehicular crimes while having little if any negative collateral damage to individuals and communities. However, viewed through a Critical Race Theory (CRT) influenced Public Health Critical Race Praxis (PHCRP) framework emphasizing racialized history and context, traffic stops have and have had clear harms at the individual, interpersonal, institutional, and cultural levels that must be weighed against any benefits to public health and safety, especially when considering disparate impacts in communities of color and low-income communities. This chapter critically examines the history and current practice of law enforcement traffic stops in the United States through the lens of PHRCP's four main foci and ten principles, offering a comparison between conventional and critical frameworks. The ten principles are also used a self-critique tool of this dissertation, identifying limitations. Through that examination, this chapter offers a model for design and interpretation of future studies and possibilities for action so that public health can both better answer calls to improved, anti-racist, activist scholarship and consider critically collaborations and conflict with law enforcement traffic stop programs.

5.2 **Background**

What are Critical Race Theory (CRT) and Public Health Critical Race Praxis (PHCRP)?

Critical Race Theory (CRT) 'defines the set of anti-racist tenets, modes of knowledge production, and strategies a group of legal scholars of color in the 1980s organized into a framework targeting the subtle and systemic ways racism currently operates above and beyond any overtly racist expressions' [46]. Further, CRT distinguishes itself from both colorblind approaches to racism, such as a feminism or class critique disconnected from intersectional race realities, and civil rights approaches that may seek redress on behalf of race without changing the underlying racist structures within those systems. Following a call for CRT incorporation in the study and teaching of education in 1998, Dr. Chandra Ford and Dr. Collins Airhihenbuwa called for its inclusion in the public health sphere in 2010 [47].

Reiterating that call in 2018 [46], Ford and Airhihenbuwa again promoted the Public Health Critical Race Praxis (PHCRP) as a semi-structured framework to facilitate the integration of CRT into public health research disciplines, such as but not limited to epidemiology, that produce and interpret empirical evidence used for intervention evaluation and policy promotion.

The PHCRP has four foci and eleven affiliated, interrelated principles. The four foci are (1) contemporary patterns of racial relations, (2) knowledge production, (3) conceptualization and measurement, and (4) action. The eleven principles that relate to one or more foci are: (1) race consciousness, (2) primacy of racialization, (3) race as a social construct, (4) gender as a social construct, (5) ordinariness of racism, (6) structural determinism, (7) social construction of knowledge, (8) critical approaches, (9) intersectionality, (10) disciplinary self-critique, and (11) voice.

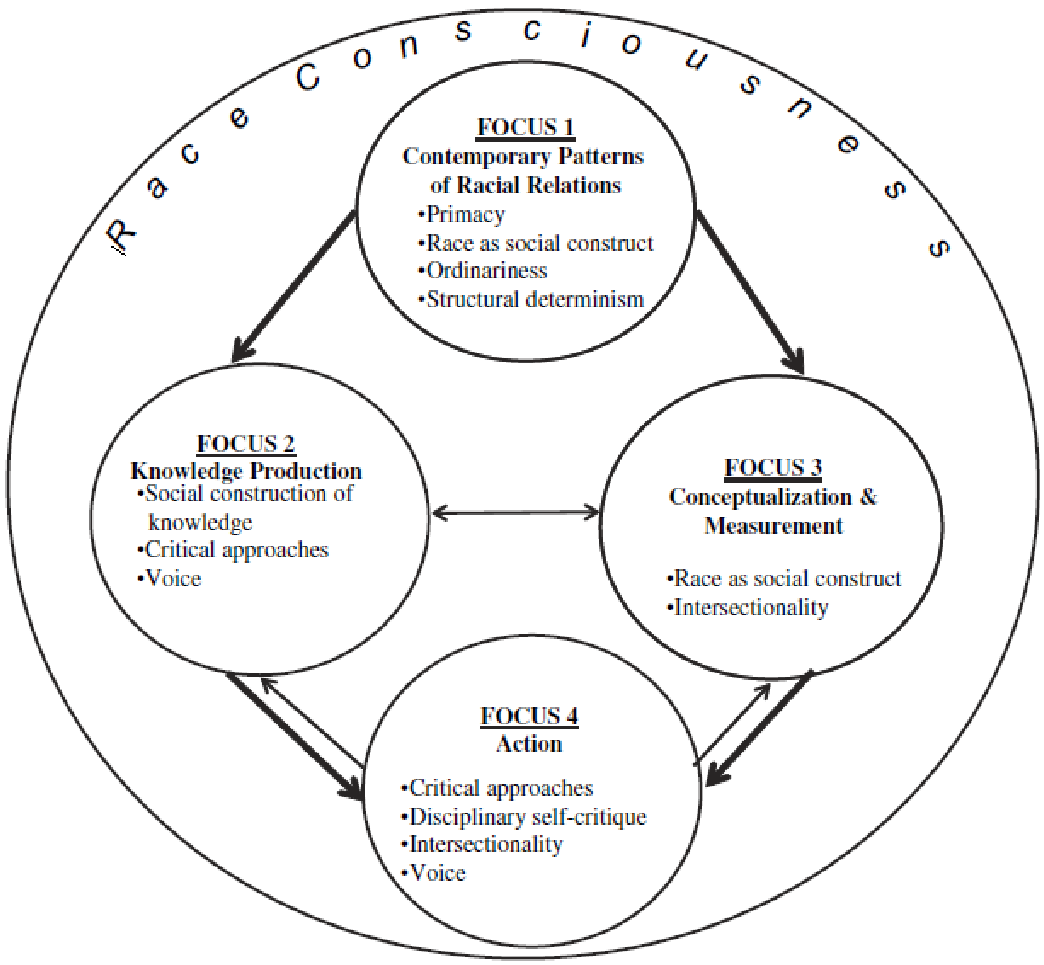


Figure 5.1 Race consciousness, the four focuses and eleven affiliated principles.

Reprinted from Ford & Aihihenbuwa, 2010.

Though applied by a decentralized network of researchers, a multi-day training institute has also been organized to provide opportunities to integrate the CRT / PHCRP principles into participant's research [46], demonstrating the possibility of integrating the framework into formal public health education programs. Since I as this dissertation's author have not been formally trained in PHCRP, I humbly offer its application here as much to demonstrate its application to traffic stops as to advance his own understanding of the framework. However, I draw on previous experience with community anti-racism training (through dismantling Racism Works and the Racial Equity Institute), both of which draw on frameworks from the People's Institute for Survival and Beyond [128]. These characteristics of institutions [103] overlap with PHCRP in many ways. See Discussion for a self-critique of this dissertation through the lenses of PHCRP and White Supremacy Culture.

Since its introduction, CRT and the PHCRP have been increasingly used to guide study design, interpretation, and suggest areas for future research. These applications are varied, recently including: a study of public park features in Latino immigrant neighborhoods [52]; interpreting the results of a survey of Black youth impacted by lead water contamination in Flint, MI [97]; and providing a guiding framework for studies of law enforcement "justifiable" homicides of Black men [52]. This chapter will follow the examples of Gilbert & Ray by contrasting a conventional interpretation with a PHCRP interpretation for each principle. Following that, the themes from the principle comparison will be combined into a figure describing these dynamics within the unique nexus of a traffic stop.

5.3 **Applying CRT / PHCRP to traffic stop frameworks**

Based on the literature review, recent community practice, current disparities, and the framing of current measurement tests, we constructed a diagram to contrast the conventional traffic stop framework to a framework informed by PHCRP. This table uses the example from Gilbert & Ray with the first three columns exactly reprinted, and the second two columns novel to traffic stop frameworks.

Principle*	Affiliated focus*	Definition*	Conventional approach	PHCRP approach
1. Race consciousness	All	Deep awareness of one's racial position; awareness of racial stratification processes operating in colorblind contexts	"Color blind" traffic stop interactions based on "objective" measures of crime and universal application of law. Race and of officer, driver, and passengers are irrelevant; demographics of neighborhoods, agencies, and political representation. Ignores existing stratification by race (e.g. segregation, income disparities, power and representation disparities, infrastructure investments) that further feed traffic stop disparities.	Understand role of individual race identities in decision making and interactions, e.g. internalized superiority and inferiority in implicitly and explicitly biasing interpersonal interactions. Acknowledge highly discretionary application of law and disconnect from measures of public health impact. Understand organization and neighborhood-level identity and demographic dynamics. Acknowledge and act equitably (not objectively) given racially asymmetrical distribution of stratification (e.g. segregation, income disparities, power and representation, infrastructure investments). Adopt actively anti-racism frameworks.
2. Primacy of racialization	Contemporary racialization	The fundamental contribution of racial stratification to societal problems; the central focus of CRT scholarship on explaining racial phenomena	Framing racial disparities as negative collateral byproducts instead of primary consequences of policing. Defensiveness on accusations of racial bias in interpersonal actions or decision making or when challenged by disparities in outcomes (e.g. differences in stop, search, etc. rates).	Acknowledge primacy of racialized policing, especially war on drugs and modern-day treatment of epidemics and poverty. Center histories of White supremacist law setting and the primary effectiveness of racism as an organizing suppression strategy. Contrast conventional frameworks with CRT frameworks for building study designs and interpreting results.

Principle*	Affiliated focus*	Definition*	Conventional approach	PHCRP approach
3. Race as a social construct	Contemporary racialization, conceptualization and measurement	Significance that derives from social, political, and historical forces	Race is only conceived as an immutable, self-identified, biological construct. Race is synonymous with phenotype. No discussion of place- and time-specific changing definitions of race, self- and other-ascription of racial identity. No discussion of strengths and limitations of categorizing diverse people's phenotypes, cultural and language experiences, self- and other-ascribed identities, ancestry, etc., in limited race-ethnicity boxes. No discussion of political forces (capitalism, White supremacy) that drive disparate treatment by race.	Acknowledge nuanced dynamics in assessing race, including place-specific passing (e.g. as White non-Hispanic), self- or other-identification of race-ethnicity, and the changing social definitions of race categories. Describe the legal treatment and protection of race and disparities juxtaposed against polices to promote White supremacy explicitly and implicitly. Contextualize traffic stop programs in decades of racism in general and law enforcement racist policies in particular: e.g. historical and present-day racialized war on drugs, enforcement of land use decisions, social control and broken-window policing.
4. Gender as a social construct	Contemporary norms of masculinity, conceptualization and measurement	Significance of gender constructions that derive from social, political, and historical forces	Ignores contemporary masculine culture norms of officers and agencies, presenting them as gender-less or gender-neutral. Ignores gender demographic dynamics in driving. Ignores the place-specific, localized construction of gender norms and demographic distributions through policy enforcement (e.g. arrest of Black men for non-violent crimes, specific driving distributions)	Names, interrogates, and may act on masculine cultures aspects of enforcement: lone wolf policing, hierarchies, officer resistance to community authority, independence, binary thinking, production and individual advancement over community relationships. Gender-specific analyses of both drivers and officers, with critical discussion of measurement. Place-level analyses that acknowledge localized gender cultures.

Principle*	Affiliated focus*	Definition*	Conventional approach	PHCRP approach
5. Ordinariness of racism	Contemporary racialization	Racism is embedded in the social fabric of society	Racism is framed as a rare event between individuals (e.g. officer and driver), instead of a multi-level, pervasive oppressive force through history that produces experiences at all levels, including micro-aggressions, explicit racial discrimination, implicit bias, institutional policies, cultural preferences, and local, state, and national policies.	Racism and its products (including traffic stop disparities) are discussed not only as (common) events, but a pervasive system that that disallows the possibility of neutral interactions or policy and demands an explicitly anti-racist approach. Focus pulls back from single opportunity for racism (e.g. individual officer bias) to multiple opportunities for individuals, agency policies, and other related content areas (e.g. driving, poverty, representation) that interact at the nexus of traffic stops.
6. Structural determinism	Contemporary racialization	The fundamental role of macro-level forces in driving and sustaining inequities across time and contexts; the tendency of dominant group members and institutions to make decisions or take actions that preserve existing power hierarchies	Sole focus of disparities is behavioral: behavior of the officer (e.g. explicit or implicit bias) and behavior of the driver (whether any behavior of could remotely, under any law, be rationale for a stop). No treatment of macro-level forces like income disparities, historical and current community disinvestment, patrol priorities or distribution. Agency and officer denial of responsibility to any structural causes in lieu of a tunnel-vision focus on whether a very specific interaction, separated from its contexts, could be rationalized. Focus on the behavioral is framed as objective, colorblind application of law and policy, even history reveals they were not constructed objectively.	Analysis of traffic stops expand beyond the immediate and behavioral to institutions (e.g. law enforcement agencies) accounts for other structural disparities and may include multi-level components. Acknowledgement of pervasiveness of structural determinism, acknowledges and moves past defensiveness to wider conception of collective responsibility (especially parts of oppression that are no one person's job, e.g. a racism without racists. White dominant institutions and white people in particular pay particular attention to disparate and compounding impacts, not just localized intentions. Institutions are directly accountable to a broad diversity of other communities and institutions, given the interrelatedness of structural determinism.

Principle*	Affiliated focus*	Definition*	Conventional approach	PHCRP approach
7. Social construction of knowledge	Knowledge production	The claim that established knowledge within a discipline can be reevaluated using antiracism modes of analysis	Data collected on traffic stop forms (including race-ethnicity and gender identifiers), associating driving data, law enforcement administrative data (e.g. court fines and fees, arrest data) are all treated as objective with known, external meanings. Little attention given to hidden dynamics or limitations data generation process. Conventional frameworks are treated uncritically as universal, immutable, and ahistorical, without an origin in time, place, people, or power.	Quantitative data, qualitiative data, and implicit and explicit frameworks that drive meaning are treated as if they have social origins and are socially mutable, especially through a power lens. This include questions like why as many traffic stops occur as they do, when did those efforts start, and how have they changed; what do traffic stops prevent, when did we come to believe this, and what evidence exists for it; what has race-ethnicity meant in the past or in different places, how does racism operate now, and how might anti-racist action operate here and now.
8. Critical approaches	Knowledge production, action	A social psychological approach to develop a comprehensive understanding of how individual biases develop prejudice and discrimination in social interaction	Knowledge produced is done so uncritically, with little attention to origin, deeper meanings, flaws, or implications. No consideration of data, information, knowledge, or wisdom hierarchy and how knowledge does or does not spread to others or deepen over time. Narratives are simple and likely separated from any considerations of shared responsibility, historical meaning, or possibility of wrong-doing on part of officers or government - excepting perhaps "bad apples" that are (again, uncritically) known to be explicitly racist.	Data, assumptions, knowledge, and actions are all examined critically, particularly with an anti-racist lens. Agencies and governments share responsibility for not just enforcing, but perpetuating racism. Anti-racist agencies continually look for places to take improved action or stop action entirely if damaging to marginalized groups. Critical voices from community members and outside agencies are not ostracized and "othered," but welcomed and integrated.

Principle*	Affiliated focus*	Definition*	Conventional approach	PHCRP approach
9. Intersectionality	Conceptualization and measurement, action	The interlocking and multiplicative approach to co-occurring social categories (e.g., race and gender) and the social structures that maintain them	Failure to consider the interacting dynamics of racism alongside sexism, homophobia, and capitalism - e.g. implicit and explicit suggestions that race and racism operate the same for all people using or ascribed a certain identity / label). Failure to adopt a multilevel approach to addressing disparities - e.g. focusing exclusively on implicit bias training and behavioral interventions.	Address white supremacist culture components alongside (toxic) masculinity cultures and other privileges and marginalized identities. Act from a multi-tiered approach when addressing disparities, considering not only personal, but institutional and cultural levels of actions, e.g. considering patrol patterns and neighboring agency practices. Integrate traffic stop program interventions alongside anti-racist public health interventions in other areas, such as overdose and mental health response.
10. Disciplinary self-critique	Action	The systematic examination by members of a discipline of its conventions and impacts on the broader society	Critical voices in local government, public health, and law enforcement are suppressed in favor of a united front. Exceptional stories and counter examples are unwelcomed. History is generally ignored, especially any history that paints a discipline in a negative light (e.g. racist history of policing, public health, and local government social control).	Critical voices are esteemed, rewarded, and developed. Critical frameworks are included in required training and treated as a conveyable skillset, not a magic alignment. History of intentional and unintentional racism within the discipline are taught with a focus on anti-racist action and change.

Principle*	Affiliated focus*	Definition*	Conventional approach	PHCRP approach
11. Voice	Knowledge production, action	Prioritizing the perspectives of marginalized persons; privileging the experiential knowledge of outsiders within	Law enforcement is the sole voice in determining programs and producing knowledge about those programs, perhaps with some minimal accountability to local government. Marginalized population experiences can be "swept under the rug" because they may be relatively few. White and middle-class experiences are taken as the overall norm, driving attention away from experiences of marginalized groups. Exceptional events are treated as necessary sacrifices to maintain otherwise effective traffic stop programs. Only law enforcement determines whether programs work, their efficiencies, and the benchmarks of success.	The stories and experiences (individually and collectively) of people who are stopped are prioritized, particularly those who are most marginalized (people of color, justice involved populations, non-English speakers, etc.). These communities lead determinations of not only how analysis is done, but how stop programs operate. In short, individuals and communities self-determine how they want to be patrolled and policed, or at least co-design stop programs with local agencies. The voice of those who are injured (e.g. by traffic crashes, assaults, or injuries from the justice system) are held up.

Table 5.1 PHCRP vs. conventional view of traffic stop frameworks.

Columns marked with (*) are reprinted verbatim from Gilbert & Ray (2015).

5.4 **Multi-level, dual-agent PHCRP framework for traffic stops**

These conventional and PHCRP principles could expand beyond this table. However, we believed there may be utility in having a more condensed figure that demonstrates the PHCRP framework more visually. The PHCRP framework when expressed tabularly does not convey (1) the nested, multi-level dynamics of people, inter-personal interactions, institutions, and cultures, and (2) does not separate drivers and residents from law enforcement as unique loci for critical analysis, with their interaction being the nexus of the traffic stop. To that end, we build the following visual framework to contrast PHCRP and conventional frameworks, nested within these multi-level structures, and separating law enforcement and driver / residents. The figure is repeated on the left and right for ease of reading small fonts.

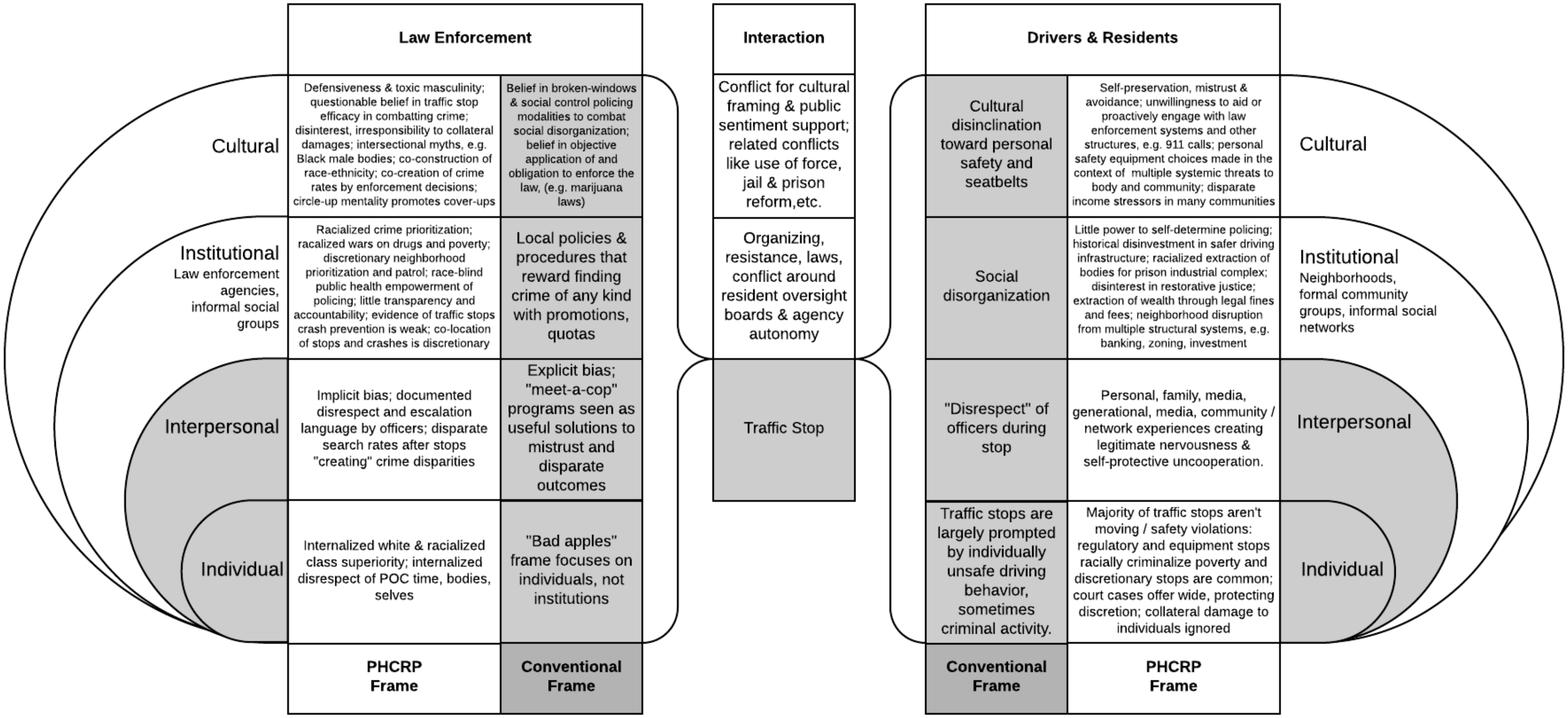


Figure 5.2 Nested, dual-agent PHCRP framework for critical examination of traffic stops.

Conventional frameworks prioritize the individual (behaviors and internalized mindsets) and interpersonal levels, and limit interaction to focus on the traffic stop itself as a time and level of interaction. PHCRP emphasizes higher levels dynamics (institutional, cultural), root and historical causes, and collateral consequences.

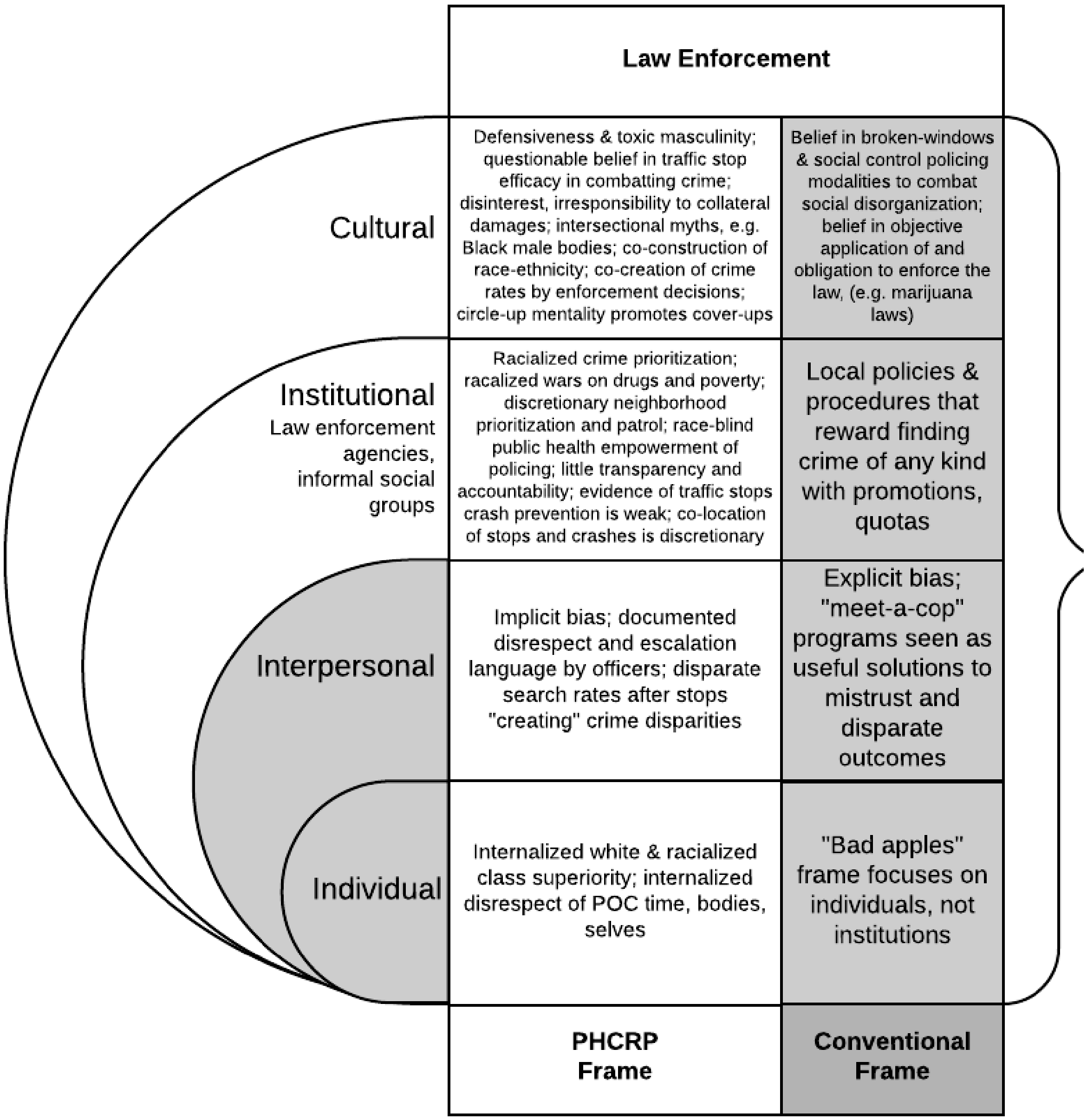


Figure 5.3 Nested PHCRP framework, left side

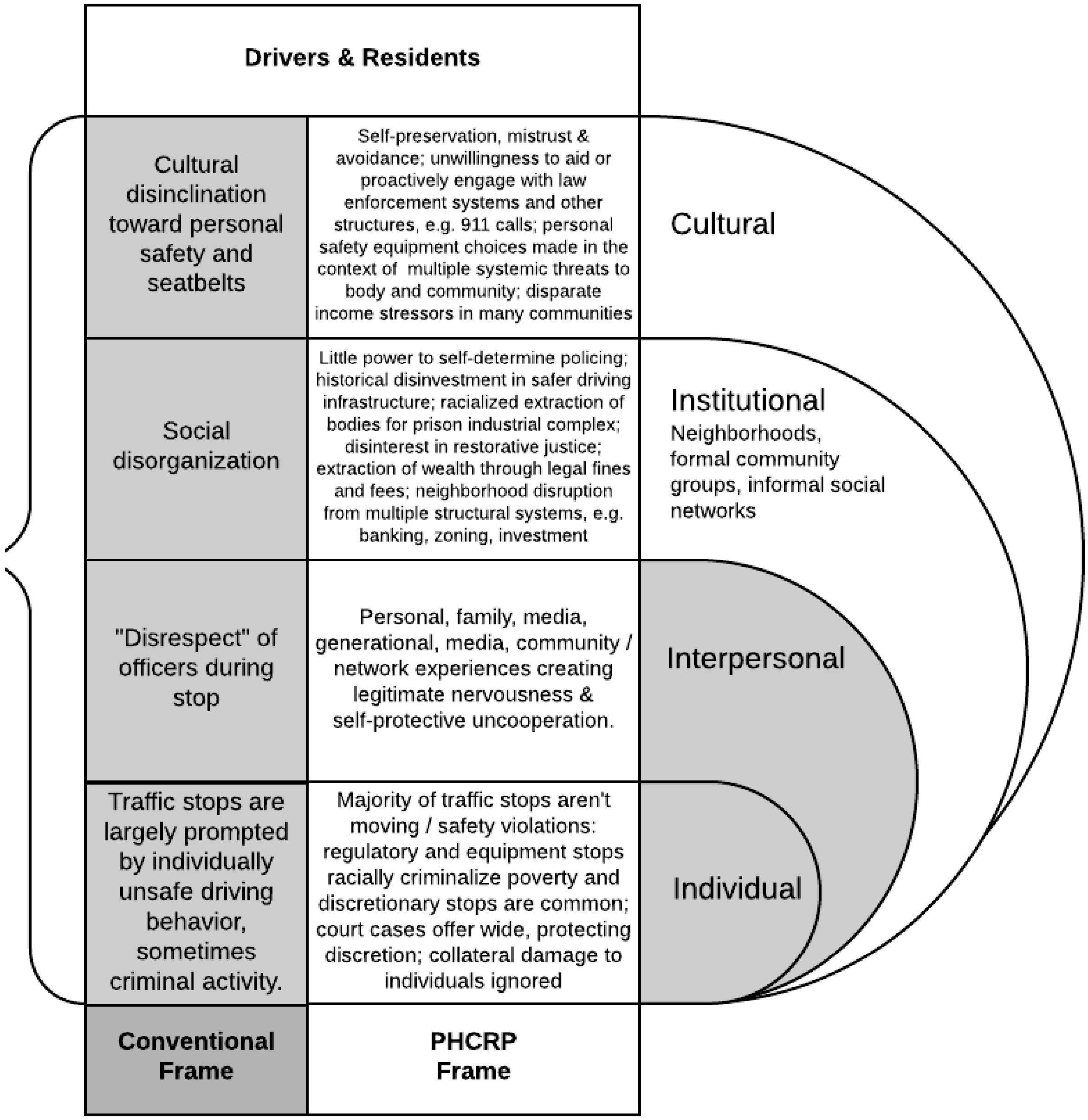


Figure 5.4 Nested PHCRP framework, right side

5.5 **Discussion**

The PH-CRP may be cautiously recommended as a framework to help guide more equitable and less unjust traffic stop policies and public health / law enforcement collaborations. However, such applications are likely to fail if not deeply applied or done without leadership from, or at least collaboration with, impacted communities. The PHCRP should not be used for rubber stamping intervention. In contrast, a truly critical framework must contend with the possibility that few to no aspects of traffic stop programs may be equitable under the PHCRP. However, it is possible that, given training in PHCRP, communities, public health, and law enforcement may co-design traffic stop programs that are tightly limited by ethics, efficient and effective in application, and serve to deepen community trust instead of endangering it. Whether this is overly idealistic or can be done in practice is yet to be determined.

The nested, dual-agent PHCRP framework captures many nuances that the tabular form does not. However, both frameworks would benefit from consideration along the time axis, currently left out of both models. Sequencing important traffic stop related moments in time (e.g. pre-stop, stop, potential citation, potential search, potential arrest), under a multi-level framework, may help identify contrasting conventional and PHCRP frameworks but also identify the various disparity considerations and tests at those moments. This is explored more in the Discussion section as an area for future research.

CHAPTER 6 - LAW ENFORCEMENT TRAFFIC STOP DISPARITY MEASUREMENT REQUIRES VEHICLE ACCESS, TRAVEL VOLUME, AND MULTI-AGENCY DRIVING CONSIDERATIONS

6.1 Overview

Introduction: Law enforcement traffic stops are one of the most common entryways to the US justice system, with significant downstream impacts for both individuals and communities. Group-specific rates are typically based on agency jurisdiction's resident populations; these rates, like many justice-system indicators, suggest race-ethnicity disparities. However, residential-based implicitly assume race-ethnicity groups have equal access to vehicles, equal annual driving volume, and that all driving occurs in resident's jurisdictions. In contrast, surveys suggest Black non-Hispanic and Hispanic households have less access to vehicles and drive less than White non-Hispanic households. This analysis reports the direction and degree of change in a disparity index when accounting for driving factors of access, driving volume, and cross-agency driving.

Methods: Data from over 20 million traffic stops in North Carolina are combined with US Census data and race-ethnicity driving factors from the 2017 National Household Travel Survey (NHTS) to calculate a disparity index based on traffic stop rate-ratios (TSRRs) under multiple model assumptions. Spatial simulation models prorate access, volume, and cross-agency driving parameters individually and together to distribute Vehicle Miles Traveled (VMT) and rebuild disparity indices for 177 law enforcement agencies.

Results: Adjusting for three driving factors simultaneously, agency disparity indices increased 15% on average from 2.02 (1.86, 2.18) to 2.33 (2.07, 2.59) for Black non-Hispanic drivers. TSRRs were largely unchanged moving from 1.43 (1.32, 1.54) to 1.38 (1.24, 1.51) for Hispanic drivers. All models suggested both groups experience disparate traffic stop rates compared to White non-Hispanic drivers.

Conclusions: Results suggest residential-based traffic stop rates may systematically underestimate already consistent disparities when driving factor differences compound. Agencies should make efforts to base traffic stop rates and disparity measures on travel-informed baselines whenever possible, though may use more simplified driving models in practice.

6.2 Introduction

Law enforcement traffic stops are the most common interaction with the law enforcement [33], serving as an entryway to the US justice system, with significant downstream and disparate impacts for individuals and communities [69]. However, states have only recently required agencies to collect and report consistently [16], even if communities have recognized these disparities for decades [59]. The though data does not exist for earlier decades to validate this, traffic stops may be increasing as a policing tactic, creating an increased need to assess disparities. Supreme court cases in 1968 and 1996 [26,80] enabled US law enforcement, under any reasonable suspicion and the loosest definitions of crime profiles, to escalate minor traffic violations into a traffic stop. Combined with the driving reality that nearly all driving includes actions interpretable as infractions [16,89], these rulings permit law enforcement nearly complete discretion over traffic stop enforcement legally, even if the public views those stops as unfair [90].

Measuring disparities

These stop rates are typically based on residential populations instead of driving populations and driving patterns that cross multiple jurisdictions. However, this approach is known to be flawed [48,132,145,148] since traffic stops are fundamentally based on vehicle driving patterns. Because of this, though preliminary data suggests already significant racial disparities in traffic stops [16], these differences may be underestimated. This analysis compares residential-based rates to more accurate, driving-based stop rates for 177 hundred law enforcement agencies, including most municipal police departments and all rural sheriff departments, using 20 million North Carolina (NC) traffic stops from 2002 to 2018, the nation's oldest and most complete traffic stop dataset [16].

Numerous individual and systemic factors combine to create differences in traffic stop counts between populations. However, it is useful to distinguish which differences are due to inaccurate or incorrect rate denominators, and once accurately constructed, which are rationale that may support or deny the existence of an unjust disparity. Herbert et al. [66] distinguish differences from disparities by the degree of agency an individual has to affect the outcome vs. structural factors like environmental and social influences. In this case, differences by race-ethnicity in stop rates by a law enforcement agency within its jurisdiction may be due to inaccurately counting the residents and their amount of driving. Disparities may be caused by factors outside of an individual's control, including institutional factors like unequal patrol patterns in neighborhoods or interpersonal implicit or explicit bias by officers. When considering differences that may constitute racial-ethnic disparities born from systems of structural disadvantage, it is useful to look beyond individual behaviors in isolation, and consider individual internalized inferiority and superiority; interpersonal bias and stereotypes, whether explicit or implicit; institutional factors such as policies and laws; and cultural effects on media and social groups. One cannot effectively consider whether differences amount to unjust disparities if the underlying differences are mis-measured.

Access, volume, and multi-agency driving

Authors have raised many potential covariates to improve interpretation of residential-based denominators in stop rates [16,48,145]. For the purpose of improved estimation of racial-ethnic differences in stop rates, we focus on three: (1) access to a vehicle, (2) total volume of annual driving, in vehicle miles traveled, and (3) drivers accumulating VMT in patrol areas of not just one, but multiple agencies. Data from the National Household Travel Survey (NHTS) can be used to better understand differences in these driving factors by race-ethnicity [87,91], and has been similarly used to better understand disparities in motor vehicle crashes [64], a phenomenon similarly connected to driving realities.

For vehicle access, known disparities in income, among other factors, lead to differences by race-ethnicity in access [20,91]. Nationwide, around one in four Black non-Hispanic, Hispanic, and Native American residents live under the poverty level, compared to one in ten White non-Hispanic residents. Consequently, previous national studies reported that more than four out of five White non-Hispanic households have access to a vehicle, compared to just half of Black and Hispanic households [126]. Notably, only access, not ownership, is required to be at risk of a traffic stop: while license or vehicle registries may be promising sources of informative data on drivers local to a jurisdiction, those without licenses and undocumented workers unable to obtain licenses also drive in order to meet basic activities of daily living.

Total volume of driving is likewise different by race-ethnicity. Type and location of jobs to commute to, the spatial spread of other activities of daily living, the cost of car maintenance and impact of income disparities, and the distribution of social amenities and networks all impact communities differently by race-ethnicity. Consequently, nationwide analysis of the 2001 NHTS suggests that White non-Hispanic households drive approximately 11,000 miles per year compared to Black non-Hispanic and Hispanic households, which each total closer to 9,000 miles [126].

Driving across multiple agency jurisdictions is commonplace: a single trip may require movement between cities patrolled by different municipal police departments, through rural areas patrolled by sheriff departments, and along highways patrolled by state highway patrol. Driving patterns typically cluster more activity nearer, and less activity farther, to core activity spaces such as homes, but these dynamics can be region specific. One study of mobility challenges for households in poverty based on the 2009 NHTS demonstrated this diversity: daily average travel radius of licenses drivers at or below poverty level was over ten miles less in Atlanta and Los Angeles, but 15 miles greater in New York City, when compared to drivers making over $100,000 a year [91].

A visual summarizes these dynamics in a simplified model of a city police department inside a county patrolled by a sheriff is in Figure 1. Access, volume of driving, and distribution of VMT into multiple agency dynamics change the rate denominator, then changing the resultant race-ethnicity-specific incident rate ratio.

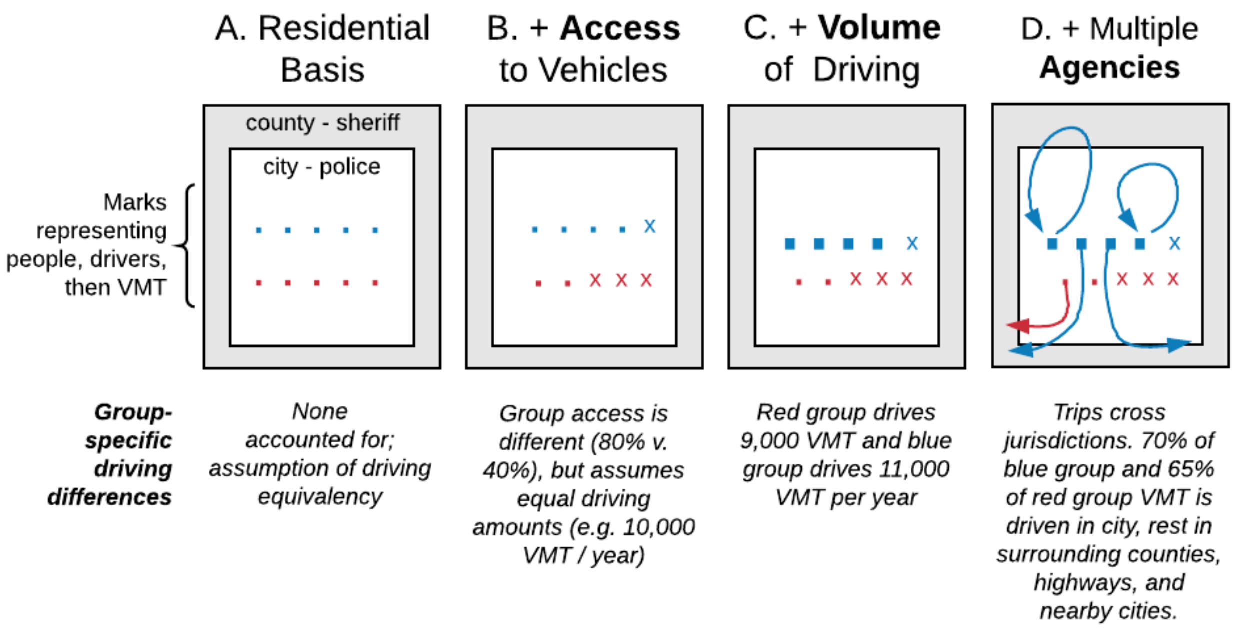


Figure 6.1 Effect of differences in driving factors between groups on disparity measures.

Other measurement models

Some authors have suggested other factors to include when considering racial differences, including agency-specific decisions to patrol different sub-jurisdiction neighborhoods differently and differences in individual driver behavior (e.g. safe movement) [48]. Fridell raises these issues but does not provide practical guidance on how to implement them. While important for interpretation of stop rates, we leave these covariates like out of the calculation of stop rates for three reasons. First, previous work separating differences from disparities suggests these value-laden constructs are better used to substantiate or defend possible unjust disparities [66]. Second, by using a rate definition consistent across agencies, agency rates can be combined and compared. Lastly, though similar behavioral and institutional factors are present in other similar rate measures, such as vehicle crash rates, they are used in interpretation and further study of those rates, not in baseline vehicle miles traveled (VMT) calculations [51,64].

Authors have proposed alternatives to residential-based traffic stop rates. Observing that vehicle crashes by race-ethnicity are derived from a similar driving distribution, Withrow and Williams [145] advocate a ratio based on at-fault vehicle crashes. This measure correctly acknowledges a shared underlying driving distribution between vehicle crashes and traffic stops, and lay understanding may suggest that traffic stops should reasonably parallel crashes if traffic stops are primarily meant to prevent vehicle crashes. However, in many jurisdictions fewer than half of traffic stops are due to moving and safety violations [16], suggesting the direct coupling of measures may be inappropriate. Additionally, because even traditional odds ratios are notoriously difficult to correctly interpret and compare [134], this measure lacks easy interpretability – a key concern for measures debated publicly by community groups and law enforcement. The same interpretability concern applies for novel techniques borrowed from motor vehicle crash literature at the sub agency level designed to address similar issues [107].

Research Triangle Institute has created an online tool (RTI STAR) based on the work of Grogger & Ridgeway [61] designed to assess racial bias in traffic stops. Acknowledging the challenges in residential denominators and survey-based approaches to answering those limitations, they use a “veil of darkness” approach that “asserts that police are less likely to know the race of a motorist before making a stop after dark than they are during daylight” [61]. They constrain analysis only to stops just before and just after sundown so the model can describe differences police behavior based on being able to identify from afar the race-ethnicity of a driver. However, this is based on a highly limited notion of potential causes of racial disproportionality – interpersonal prejudice at the time of the potential police stop by individual officers noting the race-ethnic phenotype alone (in this case, skin color) of the driver. This limited, behavioral-focused definition of disparities is not in keeping with critical anti-racism literature that describes racism as fundamentally structural and multilevel [47,54], nor crime-concentration literature advocating a multi-level approach [78], nor motor vehicle crash literature that acknowledges multi-level factors [7]. Racism and discrimination operates at reinforcing, multi-level scopes of influence: (1) internalized in an individual (as racial inferiority and / or superiority), (2) interpersonal interactions and relationships, (3) institutional (e.g. policies, laws, practices), and (4) cultural (norms, symbolism, etc.) [103]. As example of RTI STAR limitations, note that this model would fail to identify disparities if stop rates were equally high before and after sundown, even if those rates were much higher than white neighborhoods.

VMT: Following common practice

While the practice of measuring traffic stop rates is relatively new, vehicle miles traveled (VMT) have been the chosen denominator for describing driving events like crashes for decades [133]. Traffic stop rates based on vehicle miles traveled therefore have benefits in interpretability and existing measurement infrastructure. This analysis estimates the direction and degree of the mismeasurement of racial-ethnic differences in law enforcement traffic stops rates when using residential denominators instead of more appropriate driving, VMT-based denominators. I conclude with recommendations for law enforcement agencies and community groups in how to measure traffic stop rates in the field, balancing interpretability, practical considerations, and accuracy.

6.3 **Methods**

To assess the degree and direction of difference when using race-ethnicity-informed driving denominators instead of residential denominators in assessing group-specific stop rates, we use spatial simulation and driving factor estimates from supplemental datasets to derive and compare traffic stop rate ratios (TSRRs) under multiple model assumptions.

The primary dataset for analysis is the North Carolina State Bureau of Investigation (SBI)'s database of over 20 million police traffic stops from 2002 to 2018, representing 308 of the 518 state, county, municipal, campus, and place-specific (e.g. state fairgrounds, capital building) police departments [16]. By 2002, reporting was mandated by most North Carolina agencies, including all sheriff departments, state agencies, and municipal agencies above with jurisdictions about 10,000 population, making it one of the oldest and most complete traffic stop databases in the nation [16]. The population coverage by agency jurisdiction suggests this dataset includes a near-census of over 95% of the all police traffic stops of vehicles in North Carolina in this period. Basic demographic data for this dataset compared to North Carolina is included in Table 1.

	NC Traffic Stops		NC Resident Population	
	#	%	#	%
Race-Ethnicity				
White	13,258,385	58.1%	6,223,995	65.3%
Black	7,076,618	31.0%	2,019,854	21.2%
Hispanic	1,779,330	7.8%	800,120	8.4%
Native American	181,402	0.8%	108,829	1.1%
Asian	262,926	1.2%	206,579	2.2%
Other	273,176	1.2%	176,106	1.8%
Total	**22,831,837**	**100.0%**	**9,535,483**	**100.0%**

Table 6.1 Demographic comparison of NC traffic stops and NC population.

Race-ethnicity categories are mutually exclusive by including all Hispanic identified individuals in own category and all race categories are non-Hispanic. Demographic data from 2010 census. Traffic stop data from NC SBI traffic stop dataset, 2002-2017.

To address racial-ethnic differences in access to vehicles, volume of annual driving, and driving through multiple agency patrol areas, we use NC-specific statewide estimates from the 2017 National Household Travel Survey (NHTS). For access, 82% of White non-Hispanic people and 64% of Black non-Hispanic and Hispanic people had access enough to drive at all during the sample year in NC. For amount of annual driving, NHTS suggests 10,819 VMT per year for White non-Hispanic drivers, 9,775 for Black non-Hispanic drivers, and 12,434 for Hispanic drivers. These single value adjustment factors, alongside others of interest, are included in Table 2. To model travel between agencies, we use NHTS vehicle trip data to find, for example, average trip distances for White non-Hispanic drivers was 10.4 miles, Black drivers 9.7 miles, and Hispanic drivers 12.4 miles. For a more detailed distance model, race-ethnicity specific vehicle miles traveled distributions were calculated every 1-mile radius up to 400 miles. Those distributions were translated into unidirectional spatial kerning functions that, for a given distance and race-ethnicity group, describes the VMT distributed into all points at that radius.

That model was based on the log of the radius, an inflection point at 25 miles, and interaction by race-ethnicity. Associated graphs and model details are in Figure 2.

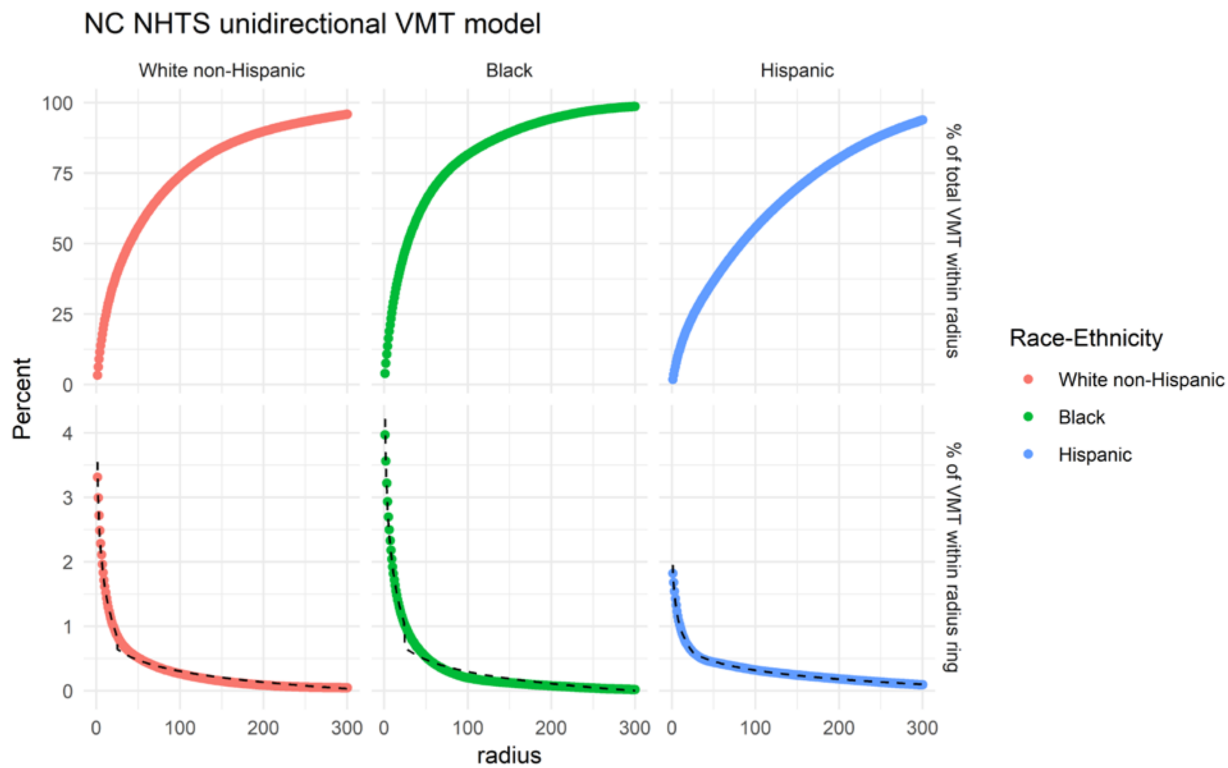


Figure 6.2 (Supplemental) Percent of ring and total VMT at given unidirectional radius.

Summary models for percent within radius (dotted black line, bottom row) were fit by the log of the radius with an inflection point at 25 miles and interaction by race-ethnicity. Hispanic North Carolina drivers drove farther on average, leading to their lower percent VMT distributed in the <1 mile radius band.

Race-Ethnicity	Measures of Survey Representation Number surveyed	Number represented	Number drivers represented
Asian	307	251,577	184,748
American Indian	156	78,171	57,496
Black	2,444	2,015,261	1,294,804
Hispanic	600	828,660	532,834
Other	522	324,620	199,508
White non-Hispanic	13,556	5,950,650	4,894,298
Total	17,585	9,448,939	7,163,689

Race-Ethnicity	Measures of Access Household has personal vehicle access (%)	Household vehicle use at least a few times a month (%)	Any driving during year* (%)
Asian	99.8	99.0	73.4
American Indian	90.3	95.4	73.6
Black	85.3	88.2	64.2
Hispanic	97.0	97.2	64.3
Other	96.1	97.6	61.5
White non-Hispanic	98.4	98.0	82.2
Total	95.8	96.2	76.8

Race-Ethnicity	Measures of Driver VMT Annual VMT per driver* (miles)	Annual VMT per person (miles)	Average miles per trip (miles)
Asian	8,677	6,372	10.0
American Indian	12,219	8,987	10.8
Black	9,775	6,280	9.7
Hispanic	12,434	7,995	12.4
Other	8,762	5,385	8.6
White non-Hispanic	10,819	8,898	10.4
Total	10,649	8,196	10.4

Table 6.2 Representativeness, access, and amount of driving by race-ethnicity in NC.

Black households have less access to vehicles, drive less often, and drive fewer total vehicle miles than White non-Hispanic drivers. Starred measures (*) were used as model adjustment factors. Data from 2017 National Household Travel Survey (NHTS).

For quantification of these dynamics, these access, amount, and multi-agency distribution estimates are transformed into parameters used to support spatial models of race-ethnicity-specific VMT distribution, traffic stop rates, and subsequent incident rate ratios. Nine models are evaluated, each adjusting zero, one, two, or all three driving factors: zero adjustment models include (1) a residential count model representing the status quo practice and (2) a driving transformed model where all residents travel the same 10,000 VMT a year; single adjustment models include (3) multi-agency driving adjustment only, (4) adjustment to amount of driving only, and (5) adjustment to vehicle access only; (6-8) double adjustment models include all pair-wise combinations of models 3, 4, and 5; and (9) a single model with all three adjustments. In all models, driving-points are uniquely assigned to patrol areas in keeping with common patrol overlap realities (e.g. sheriffs are not assumed to patrol their entire counties equally if cities are patrolled by municipal police departments). While sheriff departments may use the entire county for rate calculations, study interviews with police chiefs and sheriffs and limited supplementary GPS data suggest this adjustment is closer to the realities of patrolling.

Residents were modeled by US census 2010 counts of people attributed to census block groups, the second lowest level of spatial granularity. These residents are then prorated by access parameters into drivers, transformed by driving volume parameters into VMT estimates, then distributed over space using a unidirectional spatial density fall-off function based on the proportion of trips within each distance ring.

VMT was distributed into a 1-mile square VMT catchment raster grid (53,818 points uniformly distributed across the state) based on distance from each block group centroid to the raster point. Each point is assigned the best-match patrol area: city police departments patrol within municipal boundaries, and sheriff departments patrol county areas not patrolled by police departments. State highway patrol was not modeled in this analysis (see Discussion).

After VMT totals for each model are attributed to catchment grid points, and those point totals are aggregated into agency VMT totals, models are then standardized against a single DOT-estimated VMT total by proportionally transforming each so the total VMT for the entire system is the same 1.1 billion VMT per year regardless of model (Perdue, 2010). This standardization not only ensures model TSRR estimates are comparable but is reasonable given only one consistent VMT total was experienced by the system.

The agency-specific stop rate estimates, after modeling their rate denominators in multiple ways, are then treated as the unit of analysis to consider the direction and degree of change in the race-ethnicity-specific difference for city police and county sheriff law enforcement agencies. The distribution of agency TSRRs are combined without weighting, e.g. the distribution of IRRs is described regardless of agency jurisdiction size or number of stops of the agency. While all city and county law enforcement agency estimates were modeled, agency estimate distributions were filtered to only include 177 agencies with patrol residential populations great than 10,000, complete data over the study period (2002-2017), and at least 1,000 stops over the study period.

6.4 **Results**

Analysis of the NHTS survey to derive model parameters suggested racialized differences in access and driving amounts, specifically that Black non-Hispanic people in NC had less access to vehicles and, given access, drove less and shorter trip distances than White non-Hispanic drivers. Hispanic people had similar driving access to Black non-Hispanics, but in contrast to national studies using 2001 data [126] where their annual VMT and trip distances were similar to Black non-Hispanics, drove more and farther than White non-Hispanic drivers in North Carolina in the survey year on average.

While the subject of this study was to assess the direction and degree of change when measuring disparities in traffic stops using a travel-informed instead of residential-informed denominator, the baseline disparities in the residential-based models are noteworthy. Models 1 and 2, based on residential data without allowing for differences in driving factors, document stark disparities by race-ethnicity in the experience of Black non-Hispanic and Hispanic drivers, who were pulled over at close to twice and one-and-a-half times the rate of White non-Hispanic drivers, respectively.

After all three driving adjustments, the average agency-specific TSRR for Black non-Hispanic drivers increased 15% from 2.02 (1.86, 2.18) to 2.33 (2.07, 2.59), suggesting that using residential-based denominators alone meaningfully underestimate driving-informed rate-ratios for Black non-Hispanic residents. The TSRR for Hispanic drivers was largely unchanged, moving from 1.43 (1.32, 1.54) to 1.38 (1.24, 1.51) in the full model, a reduction of 3%.

The largest change in the estimate of the TSRR for both Black non-Hispanic and Hispanic drivers as compared to White non-Hispanic drivers was with the access adjustment, followed by the adjustment of the amount of VMT, then the multi-jurisdiction driving adjustment (see Table 3). The TSRR estimates from the single and paired model that used vehicle access and VMT amount adjustments, but ignoring cross-jurisdictional driving, returned estimates most like the full model.

		Total		Black n-H		Hispanic	
		IR (CI)		TSRR (CI)		TSRR (CI)	
Residential-based models							
M1	Residential only model	1.88	(1.59, 2.16)	2.02	(1.86, 2.18)	1.43	(1.32, 1.54)
M2	M1 scaled to total VMT	1.88	(1.59, 2.16)	2.02	(1.86, 2.18)	1.43	(1.32, 1.54)
Driving models: single adjustment							
M3	Access only	1.89	(1.60, 2.18)	2.58	(2.38, 2.78)	1.83	(1.70, 1.97)
M4	Volume only	1.89	(1.60, 2.17)	2.24	(2.06, 2.41)	1.24	(1.15, 1.34)
M5	Multi-agency only	8.85	(7.12, 10.59)	1.65	(1.46, 1.83)	1.24	(1.11, 1.36)
Two-factor adjustment models							
M6	Access & volume	1.90	(1.61, 2.19)	2.86	(2.64, 3.08)	1.59	(1.48, 1.71)
M7	Access & multi-agency	8.90	(7.15, 10.64)	2.10	(1.87, 2.34)	1.58	(1.43, 1.74)
M8	Volume & multi-agency	8.90	(7.15, 10.64)	1.82	(1.62, 2.03)	1.08	(.97, 1.18)
Three-factor adjustment model							
M9	Access, volume, & multi-agency	8.95	(7.19, 10.70)	2.33	(2.07, 2.59)	1.38	(1.24, 1.51)

Table 6.3 Simulation model results.

Adjustment of residential-based traffic stop rate ratios for race-ethnicity-specific driving factors suggest residential-based rate ratios meaningfully underestimate the greater extent to which Black non-Hispanic (n-H) and Hispanic drivers are stopped. Incident rate of residential model 1 is / 1,000 people, models 2-9 are per 1,000 VMT. All models are scaled to consistent total VMT.

6.5 Discussion

Including factors describing race-ethnicity differences in driving suggests that residential models underestimate differences of Black non-Hispanic drivers in most law enforcement agencies in NC. This is because residential models assume equal access to a vehicle, volume of driving, and driving distance. In contrast, these driving factors are different by race-ethnicity groups, leading to differences systematic underestimation in this study. These more recent NHTS-based results confirm prior literature suggesting differences in driving factors by race-ethnicity [91] and socio-economic position [126].

Model complexity

These driving models did not account for directional driving, efficient path-finding, and other driving realities that network-savvy models can better account for. However, this spatial analysis is still outside the capacity of many law enforcement agencies and community coalitions – these groups require models that maximize accuracy while compromising, where possible, on model complexity. Model complexity is driven largely by the network component of traffic models, and, in this analysis, even the unidirectional distribution of VMT using a kerning function is too computationally intensive to be used in common practice and on a regular basis by analysts in law enforcement and community groups.

The results of the nine models in this analysis suggest race-ethnicity-specific differences in access to vehicles and amount of driving were more important than modeling cross-jurisdiction driving. Differences in cross-jurisdictional driving may be the weaker of the three assumptions as well as the most difficult to accurately model for small areas. This is partly because cross-over between jurisdictions is bi-directional, e.g. not only do residents in a city patrolled by a municipal police department drive into a county patrolled more by the sheriff, but county residents likewise travel into the city. Though that travel is not equal, if it is similar then the effects of prorating each populations VMT will have minimal effect on each agency's stop rates. However, some assumptions and transformations are dependent on the choice of a difference measure, such as the TSRRs used here.

Relative vs. absolute measures of difference

Specific measures of difference are robust against specific transformations and assumptions. Multiplicative-scale measures of difference, like the TSRRs used in this analysis, are not changed if the transformation across the groups is equal on the multiplicative scale. As example, if residential populations are not constrained to an adult population old enough, and in some cases young enough, to be regularly driving, this will not have an impact on the subsequent TSRRs if the age distributions by race-ethnicity are proportionally the same. Likewise, by including those too young to drive then multiplying by group specific VMT averages appropriate for drivers, models will overestimate the total VMT of the system but not impact the underlying TSRR between groups under the same equal proportion assumption. However, traffic stop incident rate differences (IRDs) on the additive scale will be impacted by these relaxed assumptions even if group estimates are consistently skewed.

The selection of which scale to use for stop rates, e.g. multiplicative (used in this analysis) or additive, and associated measure type, relative or absolute, is not a trivial one. This decision matters not only for model integrity under different sets of assumptions, but crucially for both the definition and communication of difference magnitudes and rationale that may amount to race-ethnicity disparities. Some have argued convincingly that risk or rate differences are preferable in most cases to ratios because they appropriately scale with the underlying total incident rate and so may be of greater interest than a relative measure when considering public health significance [123,134]. Difference measures therefore capture the overall significance of the effect as well as potential modification by sub-groups.

However, measures of difference by race-ethnicity, when systemic disparities and questions of justice are considered, may have unique use for ratio measures like the TSRRs used here. In this case, consider the case of two agencies with equal total and race-ethnicity-specific populations. Agency A makes comparably few stops overall, but by systematically targeting certain neighborhoods where people of color live, it stops Black non-Hispanic drivers five times more often as White non-Hispanic drivers by VMT. Another agency stops all people five times more frequently, irrespective of race-ethnicity, but still stops Black non-Hispanic drivers 1.1 times as often by VMT rate. A ratio measure would suggest there may be more of a concern for difference in Agency A than Agency B, while a difference measure may suggest the opposite. Which is more of a concern for communities? Large ratio measures may be of unique concern for community groups concerned with equal treatment irrespective of the size of the traffic stop program. However, to appropriately characterize these relative measures, analysts should separately report a measure of the total stop rate of the agency, may find it useful to compare it to other agencies of similar size, urban make-up, crime rates, and vehicle crash rates.

These same additive vs. multiplicative relationships can be seen in the results tables. Total traffic stop rates for models including the multi-agency driving adjustment were higher (close to 9 compared to close to 2) than models that did not allow cross-agency driving, an artifact of the inclusion criteria and stabilization to a system-wide constant VMT. Smaller agencies with less stable rates were excluded from the agency rate study, but included in the simulation. Because they are smaller, these agencies systematically contribute less VMT to their larger neighbors than their larger neighbors contribute to them, creating reduced total VMT denominators in larger agencies, leading to higher total traffic stop rates for included agencies in this study. The effect on the on the traffic stop rate ratios is less pronounced, since it is also dependent on differences in the proportion of VMT contributed in all directions by race-ethnicity, not only it's total quantity.

Simpler models for practice and communication

Because of these findings, we suggest two standard models of measuring race-ethnicity-specific traffic stop rates, based on VMT, that strikes a balance of accuracy and simplicity. First, determine the common patrol area of the agency (often cities for police departments and unincorporated rural areas for sheriffs), and attribute the residential population by group to that agency. Next, using an estimate of the probability of access to a vehicle, prorate that residential population into a driving population. Drivers are then attributed an amount of VMT, and these strata specific VMTs used for the calculation of rate denominators and IRRs.

If more detailed VMT totals have been calculated by other means, as is common in major cities and for many counties by state Departments of Transportation or Motor Vehicles, this total VMTs can be divided into race-ethnicity-stratified estimates by a similar means: estimating the residential population who does the bulk of their driving in the patrol area, estimating drivers with vehicle access adjustments, estimating amount of VMT per driver by race-ethnicity, then standardizing these estimates against the better-modeled total.

For communication purposes, agencies may choose to use the overall average VMT estimates per driver to translate results back to a more person-centric measure, returning rates in the form of stops per driver per year on average. Person-centric measures may be easier for communication, though it must be acknowledged that some drivers may be stopped multiple times.

State agencies

State agencies not modeled in this analysis, such as state highway patrols or state parks, may be modeled using buffers around state highways or park areas, system wide or regional VMT estimates that are prorated as discussed previously. Prior studies suggest state agencies patrolling highways may have different traffic stop programs [16], creating outlier challenges in simultaneous comparison to sheriff and police departments, but they could be modeled similarly. For state highway patrol, since state highways are ubiquitous across most states, using the entire population of the state, even if proportionally much greater than regularly travels on state highways, will still return consistent TSRRs if the proportion of travel on state highways is similar by race-ethnicity. Other law enforcement agencies such as state parks, university campuses, or hospital police departments could be modeled similarly, by using appropriate buffers and driving-adjusted residential populations.

National models

A strength of this model that it is based on nationally available data (census products and NHTS), so this analysis could be repeated for other states. However, not all states are sampled sufficiently to provide state-specific estimates [32]. Further, sub-state, agency-specific estimates of VMT by race-ethnicity would be better than using national or even state-specific estimates. Future studies could use compare NHTS-derived factors and small-geography census data to assess whether nationally available proxies are sufficient for estimating agency-specific driving factors. The Bureau of Transportation Statistics Local Area Transportation Characteristics for

Households (LATCH) data is just such a tool, combining NHTS with census data at the tract level [22]. Though LATCH does not include race-ethnicity data, it may provide a useful basis a consistent, national method. Using a consistent method nationwide would enable law enforcement agencies and community coalitions to compare stop rates and measures of difference between agencies. Other data sources exist that may be useful to small area race-ethnicity specific VMT modeling. Division of Motor Vehicles (DMV) license data and vehicle registrations may be useful administrative datasets if race-ethnicity coding is robust, though administrative models should be adjusted to include driving by those without licenses and undocumented individuals as well. Other survey data exists, including the Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) file describing the residence and employer locations [58], though commuting is a subset of total driving. When available, individual agencies could also use local driving-based data, with rationale, to supplement their own reporting.

Neighborhood-level models

Neighborhood-level VMT denominators by race-ethnicity are required to extend these findings to more accurate assessment of stop rates within jurisdictions. Again, as above, VMT estimates could be build up by this same method, i.e. prorating residents into drivers then transforming drivers into amount of VMT, or if a total amount of VMT in an area is previously modeled, apportioning that total VMT by race-ethnicity.

Notably, a theoretical gold standard of small-area driving, e.g. exactly trip data of all vehicles by race/ethnicity of driver, likely cannot, and perhaps ethically should not, realistically be obtained in the foreseeable future. LATCH-based methods may enable improved small area VMT estimation, but neighborhood level traffic stop data is required to make use of these estimates. However, this data is still is not required by the NC form (SBI-122) and few NC agencies elect to collect spatial data on traffic stops even though GPS tools are increasingly low-cost and available. Such detailed data on traffic stops is required for more detailed assessment of Since Given theoretical models of crime concentration suggest half of all crimes occur in 4% of a city's geography [95], better small-area data on patrol activities seem to benefit police agencies. Recognizing the same, the National Institute of Justice and the Bureau of Justice Assistance collaborated with the National Household Traffic Safety Administration to promote Data-Driven Approaches to Crime and Traffic Safety (DDACTS) [31]. DDACTS includes a series of workshops, an associated journal, and techniques to formalize hot spot analysis of incidents and crashes; spatially referenced traffic stop data can not only inform prediction and intervention models, but also ensure accountability within the agency and to community priorities.

Limitations

More nuanced driving models that account for directional travel would improve the theoretical distribution of driving and may be important for sub-agency, neighborhood level driving estimation. However, model precision is only required at the unit of analysis: for agency-level measurement it is only necessary to accurately assign VMT to the correct agency jurisdiction. Further, the inclusion criteria eliminated smaller agencies and those with incomplete data, meaning the VMT the spatial model distributed into those agencies was effectively lost from the system. Given differences in driving distances, this may mean the final model underestimates Hispanic disparities and overestimates Black non-Hispanic disparities.

This analysis leaves out other race-ethnicity groups and sub-groups, including Indigenous / Native Americans and specific Asian-American groups, which make up 4% and 1% of North Carolina's population respectively. Small numbers create additional challenges in studying the systematic effects of driving denominators statewide, but in practice agencies should tune disparity analyses like these to the communities within and nearby their jurisdictions. Given income disparities in Native American and some Asian-American groups, these study results suggesting systematic underestimation may apply. In addition, previous studies have suggested that some traffic stop disparities, such as subsequent searches, are modified not only by race-ethnicity, but gender and age [16], which this analysis does not account for. This may partly be due to driving differences by gender within race-ethnicity groups [20,51]. However, in contrast to search outcomes that occur after officers view drivers face to face, traffic stop disparities may be more linked to neighborhood segregation and patrol decisions than interpersonal interactions where gender would be ascribed and implicit biases acted on. If spatial neighborhood segregation occurs more by race-ethnicity and income than gender, these dynamics may be less at play. However, disproportionate incarceration of Black men and disproportionate driving by Hispanic males are important to consider in future gender-specific analyses.

Conclusion

Simulations suggest the standard practice of using residential-based denominators for traffic stop rates may systematically underestimate race-ethnic differences if differences in vehicle access, volume of driving, and driving patterns combine in the similar directions. Nationwide disparities in socio-economic position by race-ethnicity suggest this finding may extend to agencies nationwide, but local patterns of driving (such as Hispanic drivers in this analysis) may moderate that underestimation. Importantly, under or overestimation aside, all models, residential or driving-based, demonstrated some disparity in traffic stops by race ethnicity in NC police and sheriff agencies. Instead of residential-based rates, researchers studying traffic stops by race-ethnicity should attempt to adjust for driving factors when possible, as was done in this analysis. Though not included in this analysis, by the same reasoning this guidance would extend to analysis of traffic stop disparities by other sub-groups, such as income strata, whose residential relationships that are similarly confounded by driving factors. Agencies should make efforts to base traffic stop rates and disparity measures on travel-informed baselines whenever possible, though may use more simplified driving models in practice.

CHAPTER 7 - RE-PRIORITIZING TRAFFIC STOPS FOR PUBLIC HEALTH: AN INTERVENTION IN FAYETTEVILLE, NORTH CAROLINA

7.1 **Overview**

Law enforcement traffic stops are one of the most common entryways to the US justice system. Conventional frameworks suggest traffic stops promote public safety by reducing dangerous driving practices and non-vehicular crime. Law enforcement agencies have wide latitude in enforcement, including prioritization of stop types: (1) safety (e.g. moving violation) stops, (2) investigatory stops, or (3) economic (regulatory and equipment) stops. In order to prevent traffic crash fatalities and reduce racial disparities, the police department of Fayetteville, North Carolina significantly re-prioritized safety stops. Annual traffic stop, motor vehicle crash, and crime data from 2002 to 2016 were combined to examine intervention (2013-2016) effects. Fayetteville was compared against synthetic control agencies built from 8 similar North Carolina agencies by weighted matching on pre-intervention period trends and comparison against post-intervention trends.

On average over the intervention period as compared to synthetic controls, Fayetteville increased both the number of safety stops +121% (95% confidence interval +17%, +318%) and the relative proportion of safety stops (+47%). Traffic crash and injury outcomes were reduced, including traffic fatalities -28% (-64%, +43%), injurious crashes -23% (-49%, +16%), and total crashes -13% (-48%, +21%). Disparity measures were reduced, including Black percent of traffic stops -7% (-9%, -5%) and Black vs. White traffic stop rate ratio -21% (-29%, -13%). In contrast to the Ferguson Effect hypothesis, the relative de-prioritization of investigatory stops was not associated with an increase in non-traffic crime outcomes, which were reduced or unchanged, including index crimes -10% (-25%, -8%) and violent crimes -2% (-33%, -43%). Confidence intervals were estimated using a different technique and, given small samples, may be asymmetrical.

The re-prioritization of traffic stop types by law enforcement agencies may have positive public health consequences both for motor vehicle injury and racial disparity outcomes while having little impact on non-traffic crime.

7.2 **Introduction**

Law enforcement traffic stops are one of the most common entryways to the US justice system [33]. Community-led movements [10], national press [120], peer-reviewed research [15] and the Department of Justice [132] have all suggested that traffic stops are most burdensome to low-income and racial-ethnic minority drivers and their communities. In this paper we provide a brief background on law enforcement traffic stops through conventional and critical public health lens and evaluate an intervention designed to reduce racial-ethnic disparities in traffic stops while reducing traffic crash injury outcomes.

Conventional frameworks suggest traffic stops promote public safety by reducing dangerous driving practices and non-vehicular crimes. Assumptions of criminal justice deterrence theory [17] underlie these conventional frameworks, treating dangerous driving and non-vehicular crimes as events where each actor rationally weighs the certainty of being caught, the celerity (speed) of that consequence, and the severity of punishment against the immediate positive consequences of their action. This frame suggests a seemingly objective world of traffic stop rationale where some have chosen to break the law, others have not, and traffic stops of all kinds have a wholly positive effect on public safety. These conventional frameworks either ignore traffic stop disparities entirely or justify them as negative collateral consequences to otherwise legal and rationale public safety interventions. In either case, conventional frameworks suggest these disparities merit little attention and action under an objective enforcement of the law.

Law Enforcement Discretion, Priorities, and Disparities

In contrast to conventional frameworks, public health authorities have called for analyses that center disparities like these and for engagement in anti-racist action [74]. The American Public Health Association (APHA) recently launched a National Campaign Against Racism [75]. That campaign suggests public health advocates interested in disparities go beyond an individual focus (e.g. who is or isn't racist) to ask, "how is racism operating here?" within structures, policies, practices, norms and values [75].

One mechanism for how racism operates in the application of justice is through individual and agency discretion. In contrast to the conventional frameworks emphasizing objectivity, law enforcement agencies have wide, subjective latitude in the selective enforcement of traffic stops in practice. Supreme court cases in 1968 and 1996 [26,80] enabled US law enforcement, under any reasonable suspicion and the loosest definitions of crime profiles, to escalate any traffic violation, however minor, into a traffic stop [89]. When combined with the driving reality that nearly all driving trips include actions interpretable as infractions, whether small wavering within lanes or movement over or under posted speed limits [16,89], these rulings permit law enforcement nearly complete discretion over traffic stop enforcement legally, even if the public views those stops as unfair [90].

These enforcement and patrol priorities differentially expose populations to different patrol densities and thresholds of interaction based on neighborhood-level factors. Neighborhood-level segregation by race-ethnicity and income, when coupled with institutional policies prioritizing certain spaces and incidents operate alongside any additional disparities caused by interpersonal bias based on perceived race-ethnicity phenotypes. Indeed, previous studies have quantitatively refuted the idea that individual outlier officers (e.g. the "bad apple" hypothesis) sufficiently explain the large racial-ethnic disparities found in traffic stop metrics [16]. Still, all individual officers exercise subjective discretion in their traffic stop enforcement, and all do so partly informed by their race-ethnicity, gender, and socio-economic position personal biases, both implicit and explicit. In addition, individual officers do not operate within a vacuum. Officers operate within, or at least influenced by, the implicit norms and explicit policies of their agencies. Those formal and informal policies include neighborhood-specific patrol deployments and the relative emphasis of public safety and control priorities.

The Public Health Critical Race Praxis (PHCRP), based on Critical Race Theory [110] provides a standardized framework to investigate these traffic stop dynamics. [46,47]. Applications of PHCRP often contrast a conventional framework with one informed by PHCRP's principles [97]. PHCRP principles also explicitly acknowledge the social construction of knowledge, structural determinism, critical analysis, and disciplinary self-critique [47]. In keeping with these principles, and in contrast to the conventional framework, we recognize that a law enforcement agency's priorities and exercise of discretion are constructed over time, malleable in the present and future, influence officers and communities beyond individual interactions, and deserving of critical analysis.

Considering the relative and absolute frequency of traffic stops by the type of stop is one method of understanding an agency's implicit and explicit priorities. For the purpose of this discussion, we divide traffic stops into three categories: (1) "safety stops" including violations of speed limits, stop lights, driving while impaired, and safe movement; (2) "investigatory stops" including explicit investigation, unspecified rationales, and discretionary seatbelt enforcement (that in prior studies are most similar to investigatory stops in disparate application (Baumgartner 2019); and (3) "economic stops" that are disproportionately consequences of economic circumstances, including not carrying insurance, expired motor vehicle registrations, or equipment malfunctions. Under conventional frameworks these three stop types may be associated with public safety injury and crime outcomes. For instance, safety stops ostensibly reduce motor vehicle and pedestrian crashes. Similarly, investigatory stops may be designed to reduce non-traffic crime or discover and detain individuals after having committed certain crimes. Finally, economic stops could be framed conventionally as reducing traffic crashes because of equipment failures. Because of their link to public safety outcomes, the relative and absolute frequency of these traffic stop types represent a set of often implicit public health prioritizations.

However, disparities in traffic stops may also differ by these stop types: For instance, Black and Hispanic drivers constitute a larger proportion of investigatory and economic stops than safety related stops in the North Carolina, and are disproportionately over-represented in all stop types relative to the North Carolina population [15]. In contrast with conventional frameworks that conceive economic stops and protective and unbiased, critical intersectional frameworks acknowledge the link between race-ethnicity and income disparities. Since Black and Hispanic individuals are often disproportionally represented in low-income and low-wealth populations, they may also be disproportionally at risk of economic stops individuals and, due to segregation, more likely to live in lower-resourced areas where investigatory stops are more prevalent. Further, these higher-disparity stops are not infrequent: statewide, previous analysis of the North Carolina traffic stop dataset statewide [16] demonstrates that economic and investigatory stops make up nearly half of all traffic stops. These disparities by traffic stop type suggest that an agency's relative traffic stop type priorities, whether implicit or explicit, represent not only prioritizations of public safety outcomes but also potentially disparate population targeting.

When agency and officer enforcement priorities differ from community priorities, this violates principles of community self-determination and consequently threatens community trust and perceived legitimacy of law enforcement [45,63]. Trust may also be challenged within agencies, such as when new agency priorities differ from individual officer priorities [85]. Law enforcement agencies or individual officers may respond to community mistrust and calls for increased community accountability by scaling back their public safety services (such as certain traffic stops) believed to be essential for violent crime control. This dynamic, named the Ferguson Effect [62], is therefore observable (and testable) in two parts: after increased public scrutiny or reprioritization of public safety activities, a (1) drop in activities and (2) an increase in the negative outcomes (e.g. violent crime) those activities were meant to protect against. Studies have shown evidence of Ferguson Effects in the attitudes and actions of officers (drops in productivity, reduced motivation, belief crime will rise as officers "de-police"), though this effect was moderated by their belief in whether communities afford legitimacy to policing [101]. In contrast, the evidence for increases in negative crime outcomes after de-policing is mixed, confounded by income inequality and racial segregation [62], and a recent Missouri study found no effect at all when traffic stops, searches, and arrests are reduced specifically [121]. This analysis considered just such a reprioritization within an agency after community members challenged police legitimacy, so we acknowledge this Ferguson Effect as a relevant dynamic for consideration.

Fayetteville intervention

Given finite budget and staffing realities, law enforcement administrators may choose to direct agency traffic stop programs to target certain public safety outcomes by prioritizing traffic stops by type or directing patrol patterns to maximize traffic stop efficiency. In keeping with this opportunity, city leaders in Fayetteville, North Carolina were called to respond to the city's consistently high motor vehicle crash rate [43]. Simultaneously, tensions between community groups and police came to a head as city council intervened to halt searches that disproportionately targeted Black residents. Soon after, the police chief and second-in-command stepped down [129].

After newly being appointed in 2013 and faced with issues of motor vehicle crashes and eroded community trust, Chief Harold Medlock voluntarily requested a review of his department practices and policies by the US Department of Justice Office of Community Oriented Policing Services' (COPS Office) [30] Collaborative Reform Initiative for Technical Assistance (CRI-TA) [115]. That report provided preliminary evidence of racial disparities in traffic stops compared to Fayetteville's residential data, though also documented the beginnings of a reduction starting with his tenure in 2013. The report also documented that Fayetteville newly elected to require officers collect Global Positioning System (GPS) data on all traffic stops, an element still not required on the state form; this is corroborated in Fayetteville's written policies for traffic stops, where failure to record this data are grounds for negative performance review [42]. Those data could then be used alongside its Crash Analysis Reduction Strategy (CARS) program, where ten intersections with the most crashes were used for targeted traffic stop enforcement each week [35]. To increase transparency and accountability, press releases were disseminated each week detailing these locations, with three intersections targeted each day. The press releases also detailed the written warnings and state citations issued the prior week.

Because of Chief Medlock's focus on traffic crash reductions and improving community trust exacerbated by racial disparities in traffic stops and other outcomes, he gave guidance to highly prioritize safety stops in order to prevent traffic crash fatalities and reduce racial disparities during his tenure from 2013 to 2016 [41]. We hereafter refer to this collection of changes to agency traffic stop activities, associated policies, workflows, staffing changes, and required organizational change work as the Fayetteville intervention. Notably this intervention included mechanisms we are not measuring in this analysis, including both quantifiable changes (e.g. possible increased spatial clustering of safety traffic stops around high crash locations) and changes more difficult to quantify, such as gradually changing internal organization culture and norms. Therefore, though we track four quantitative measures describing their traffic stop prioritization profile to gauge intervention implementation over the study period, they are best seen as representative indicators of the intervention, not the full substance or mechanism of the intervention.

The purpose of this paper was to evaluate this Fayetteville intervention alongside a broader examination of the relationship of law enforcement traffic stops and public health outcomes.

7.3 **Methods**

The intervention impact was assessed by comparing traffic stop, motor vehicle crash, and crime measures from Fayetteville Police Department to a composite control agency built by a weighted combination of data from eight similarly large North Carolina police departments that did not enact Fayetteville's reprioritization intervention.

Four domain areas were chosen to assess the intervention's impact. Traffic stop prioritization profile measures were chosen to provide evidence the intervention was not only designed and publicized but implemented. Traffic stop disparity measures were chosen to assess questions of improved equity. Motor vehicle crash measures were chosen to assess crashes averted and lives saved. Crime measures were chosen in order to explore the possibility of a Ferguson Effect, the possibility that a de-prioritization of investigatory and economic stops was associated with an increase in crime.

Thirteen measures were chosen from those four domain areas to assess these questions in more detail. Traffic stop prioritization profile measures included (1) number of safety-related traffic stops, (2) percent of safety-related stops, (3) percent of regulatory and equipment stops, and (4) percent of investigatory stops. Measures of traffic stop disparities included (4) percent Black non-Hispanic stops and (5) the traffic stop rate ratio (TSRR) of Black non-Hispanic to White non-Hispanic stops. Motor vehicle crash measures included (6) total crashes, (7) crashes with injuries, and (8) crash-related fatalities. Lastly, crime-related measures included violent crime (9) counts and (10) rates and index crime (11) counts and (12) rates. Notably, Black non-Hispanic traffic stop disparities against White non-Hispanic referent, though only one of a number of useful disparities to consider by race, ethnicity, gender, and age (Baumgartner et al., 2018), were chosen because of previously documented disparities, the specific history of anti-Black racism in the United States, and the explicit focus in Fayetteville around those disparities.

When considering causal questions involving race-ethnicity, individual race-ethnicity comes to simultaneously represent a range of interrelated, but separate constructs (e.g. phenotype, self-identified race, socially assigned race, experiences of discrimination, structural racism, historical trauma, etc.) that have unique causal relationships [135]. We acknowledge this, but do not in this study divide the construct into its many components or bring in accessory datasets to improve its contextualization and construct precision.

Data sources

Traffic stop data were obtained from the North Carolina State Bureau of Investigation (SBI) database, including over 20 million police traffic stops from 2002 to 2018, representing 308 of the 518 state, county, municipal, campus, and place-specific (e.g. state fairgrounds, capital building) police departments [99]. By 2002, reporting was mandated by most North Carolina agencies, including all sheriff departments, state agencies, and municipal agencies above with jurisdictions above 10,000 population, making it one of the oldest and most complete traffic stop databases in the nation [16]. Though it does not include all agencies, it represents the policing jurisdictions of 99% of the state population, excluding only the smallest cities and place-specific agencies. All traffic stop measures are available from the SBI dataset alone except for one were derived solely from this dataset.

One evaluation measure, the rate ratio of Black non-Hispanic vs. White non-Hispanic driver traffic stops, required accessory datasets to calculate. Per previous literature [92,132,145], commonly used, residential-based rates for traffic stops are fundamentally flawed since traffic stops are inherently tied to travel patterns. A supplemental dataset, the 2017 National Household Travel Survey, was used previously to produce NC-specific estimates of vehicle access and vehicle miles traveled by race-ethnicity group [92]. Since NC elected to additionally fund the survey as an add-on partner for supplemental sampling [32], survey results could be made representative of the state by multiplying by the pre-calculated weight factors specific to households, people, or trips to account for nuanced sampling strategies and non-response adjustments. Statewide estimates of vehicle access and total annual VMT (see Supplemental Table 2) were used as an adjustment factor to city- and year-specific residential data to derive city-year-specific estimates of drivers and total VMT by race-ethnicity [92]. Specifically, 64.2% of Black non-Hispanic residents of Fayetteville were estimated to have access to a vehicle, contributing approximately 9,775 VMT per year per driver on average. These driving adjustment factors were 82.2% and 10,819 VMT for White non-Hispanics, respectively.

Population demographic data for race-ethnicity-specific rate calculations were obtained from the United States American Communities Survey (ACS) and United States census, interpolating years 2002 to 2009 using 2000 and 2010 census data when ACS estimates were unavailable. Data on North Carolina motor vehicle crashes since 2002 were obtained from the University of North Carolina Highway Safety Research Center (HSRC) [131], and data on North Carolina crime counts and rates since 2002 were also obtained from the North Carolina SBI [99].

Synthetic control

Authors have recently advocated for synthetic control's utility to epidemiology [109] and it has been used specifically in assessing policy effects in both justice [55,97] and public health [4] contexts. In contrast to difference-in-difference (DiD) modeling, which can be conceived of a special case of synthetic control [146], the synthetic control techniques compare measures from one or more intervention units over time (in this case, Fayetteville Police Department is the single unit) against measures derived from the weighted combination of 1 or more units from a pool of control units [4]. Synthetic control therefore has benefits over DiD in maximizing similarity to controls, loosening the parallel trends assumption, and a statistical basis for control selection [114].

In this study, Fayetteville Police Department was the single intervention unit and eight similarly large cities in North Carolina served as the pool of potential controls (see Table 1). In this case and with small intervention (N=1) and potential control pools numbers, the synthetic control technique finds 1 or more control agencies that, in linear weighted combination, generate a synthetic agency for each outcome measure with a pre-intervention trend that maximizes similarity against the intervention agency (or units, in larger studies) on for each measure. These same linear combinations of agency weights, determined by the pre-intervention period (2002-2012) matching, are then applied to the same agencies in the post-intervention period (2013-2016). The intervention agency can then be compared to the synthetic control agencies for each measure to compared to generate an estimator of the difference between the Fayetteville with the intervention applied and a counterfactual Fayetteville as if it did not receive the intervention. Synthetic control methods, as a method of weighted matching, have the benefit of controlling for some unmeasured confounders [4,55] and can optionally be matched on one or more known time-varying or time-unvarying confounders, though this was not done here. See Table 1 for the list of cities and summary measures from the pre-intervention period.

	Demographic Measures			Traffic Stop Measures				Crash Measures			Crime Measures			
	Population	**% Black**	**Median household income**	**Average annual safety stops**	**Safety stops (%)**	**Black driver stops (%)**	**Traffic stop rate ratio***	**All crashes**	**Crashes with injuries**	**Fatalities from crashes**	**Index crimes**	**Index crime rate**	**Violent crime count**	**Violent crime rate**
Intervention City														
Fayetteville	203,670	41%	$43,882	13,968	43.8	56.8	2.5	5,298	1,886	62	13,367	7,848.1	1,224	730.5
Control Cities														
Cary	155,822	8%	$94,617	9,179	56.5	18.3	3.8	2,355	615	9	2,145	1,663.8	115	88.9
Charlotte	808,834	35%	$55,599	47,177	43.4	50.4	2.7	22,943	8,241	168	45,840	6,219.8	6,243	845.2
Durham	251,761	39%	$52,115	9,329	48.7	57.0	2.8	7,284	1,979	38	13,233	6,121.4	1,758	806.2
Greensboro	282,177	41%	$42,802	21,043	55.6	50.9	2.1	7,374	2,930	53	14,873	5,976.1	1,767	708.4
High Point	108,982	33%	$43,322	9,919	47.9	40.8	1.9	2,327	908	23	5,719	5,805.5	653	659.8
Raleigh	441,326	28%	$58,641	26,374	44.6	45.0	2.9	13,675	3,608	80	14,687	4,063.9	1,914	530.8
Wilmington	113,724	18%	$43,855	6,674	52.6	25.7	1.9	3,454	1,298	32	6,679	6,707.7	774	773.5
Winston-Salem	238,474	34%	$40,898	13,616	46.1	45.0	2.1	5,811	1,798	42	15,026	7,004.1	1,690	786.6

Table 7.1 Fayetteville and control agency demographics, traffic stops, crashes, and crime.

**Traffic stop rate ratio is White non-Hispanic to Black non-Hispanic drivers adjusted to travel denominators instead of residential denominators. Average annual data from pre-intervention period (2002-2012). Abbreviations: MHHI = Median household income.*

Race-Ethnicity	Measures of Survey Representation Number surveyed	Number represented	Number drivers represented
Asian	307	251,577	184,748
American Indian / Alaskan Native	156	78,171	57,496
Black / African American	2,444	2,015,261	1,294,804
Hispanic	600	828,660	532,834
Other	522	324,620	199,508
White non-Hispanic	13,556	5,950,650	4,894,298
Total	17,585	9,448,939	7,163,689

Race-Ethnicity	Measures of Access Household has personal vehicle access (%)	Household vehicle use at least a few times a month (%)	Any driving during year (%)
Asian	99.8	99.0	73.4
American Indian / Alaskan Native	90.3	95.4	73.6
Black / African American	85.3	88.2	64.2
Hispanic	97.0	97.2	64.3
Other	96.1	97.6	61.5
White non-Hispanic	98.4	98.0	82.2
Total	95.8	96.2	76.8

Race-Ethnicity	Measures of Driver VMT Annual VMT per driver (miles)	Annual VMT per person (miles)	Average miles per trip (miles)
Asian	8,677	6,372	10.0
American Indian / Alaskan Native	12,219	8,987	10.8
Black / African American	9,775	6,280	9.7
Hispanic	12,434	7,995	12.4
Other	8,762	5,385	8.6
White non-Hispanic	10,819	8,898	10.4
Total	10,649	8,196	10.4

Table 7.2 (Supplemental) NC representativeness, access, and volume by race-ethnicity.

*Data for North Carolina from 2017 National Household Travel Survey (NHTS). Black households have less access to vehicles, drive less often, and drive fewer total vehicle miles than White non-Hispanic drivers. Measures marked with a * were used in adjusting residential counts to approximate vehicle miles traveled for rate calculations. Reprinted from Fliss, 2019.*

In this case, the synthetic control method was chosen to control for known global time trends (e.g. statewide changes in driving frequency) that a single-unit difference-in-difference analysis would have left uncontrolled for. As example, driving frequency may have changed statewide over the intervention period as a function of changes in employment due to the recession and its recovery. Comparing Fayetteville's pre-intervention trend to only its own post-intervention trend would erroneously conflate any reduction in crashes of Fayetteville's intervention to the reduction in crashes due to global changes in statewide driving. Synthetic control provides some control of this kind of confounding. Because the specific causal relationships of the intervention and its covariates are largely unmapped and because of the relatively small number of observations (acknowledging an intervention n = 1), no attempt was made to control for other specific time-varying or time-unvarying confounders between agencies beyond confounding control that weighted matching on pre-intervention period provides for these global confounders. Independent synthetic control agencies were created for each measure for this same reason; simultaneous matching against all measures implies shared confounders between them, which was not known (and was not expected by authors) to be the case.

The post-intervention synthetic control annual average, annual difference between intervention and control, percent change with confidence interval, permutation p-value (calculated by assigning intervention status to each control agency and recalculating the post-intervention model), and linear trend p-value were calculated for each reprioritization, crash, disparity, and crime measure. 95% confidence intervals were estimated using Taylor series linearization as having relatively few units limit resampling- and permutation-based methods. Given the number of units, these point estimates may not exactly match those derived from the synthetic control weighting-based method and therefore confidence intervals may be unsymmetrical. The statistical package R [108] and key libraries [105,114,143] were used for analysis.

7.4 **Results**

Synthetic control generated measure-specific weight vectors using between 1 and 5 control agencies (see Supplementary Table 1), with the model average of 3.0 agencies. Table 2 presents annual averages, differences, and percent change comparing post-intervention Fayetteville to the post-intervention control agency for thirteen intervention-related measures. At the end of the intervention period over 80% of Fayetteville's traffic stops were safety stops, up from a low of 30% in 2010. The Fayetteville intervention was associated with a 47% average increase in the proportion of safety stops and a striking 121.3% (17.3%, 318.1%) average increase in the number of safety stops. From a low of just over 9,000 safety stops in 2006, at the end of the intervention period Fayetteville completed nearly 60,000 safety stops in 2016.

Both measures of Black non-Hispanic traffic stop disparities were reduced in Fayetteville as compared to the synthetic control agencies: the percent of traffic stops reduced 7.0% and the driving-adjusted traffic stop rate ratio was reduced 21%. Linearization estimates were similar and associated confidence intervals were relatively small.

All three measures of negative traffic outcomes were also reduced relative to synthetic controls: total crashes were reduced 13% (765 fewer each year), injurious crashes were reduced 23% (479 fewer each year), and traffic fatalities were reduced 28% (representing 19 fewer fatalities each year). The percent change in metrics associated with motor vehicle crashes were large but had wider confidence intervals and moderate agreement with Taylor linearization estimates.

Non-traffic crime outcomes showed little change. Index crime counts and rates were reduced 10% and 5% respectively, though confidence intervals were high. The Fayetteville violent crime count and rate were effectively indistinguishable from the control, with small estimates, wide relative confidence intervals, permutation test p-value > 0.99 and linear p-test of 0.96. Because of this, synthetic control estimates poorly matched the Taylor linearization estimates and small counts and rates disagreed in direction of association.

Figure 1 shows the trend of nine of these measures. The respective synthetic control agencies closely matched Fayetteville's pre-intervention trends for most measures. Relatively small numbers of traffic fatalities among many agencies created more variation in the pre-intervention match for that measure. Divergence in the intervention period (in grey) demonstrates the intervention's modeled effect.

	Cary	Charlotte	Durham	Greensboro	High Point	Raleigh	Wilmington	Winston-Salem
Traffic Stop Profile								
Total Safety Stops	-	4	75	-	-	21	-	-
% Safety Stops	-	7	-	-	-	93	-	-
% Regulatory & Equip. Stops	-	17	-	-	-	65	-	18
% Discretionary	31	58	-	7	4	-	-	-
Measures of Traffic Stop Disparity								
% Black non-Hispanic Stops	-	-	100	-	-	-	-	-
Black non-Hispanic TSRR	2	59	12	-	0	-	27	-
Motor Vehicle Crash Outcomes								
Crashes (all)	40	-	-	-	-	13	-	46
Crashes (w/ injuries)	34	-	-	-	-	31	-	35
Traffic Fatalities	26	31	-	-	43	-	-	-
Crime Outcomes								
Violent Crimes	29	-	-	-	-	3	-	67
Violent Crime Rate (/1,000)	14	49	-	-	-	11	-	26
Index Crimes	14	-	-	-	-	-	-	86
Index Crime Rate (/1,000)	-	15	-	-	-	-	17	68

Table 7.3 (Supplemental) Synthetic control weight vectors for each measure.

Synthetic controls were programmatically determined by maximizing the match on pre-intervention trends for each measure, producing weight vectors of between one and five (mean 3.0) other NC city police departments linearly combined to model post-intervention counterfactual trends.

	Fayetteville Police Department		Synthetic Control	Difference between Fayetteville and Synthetic Control				
	Pre-intervention annual average	Post-intervention annual average	Post-intervention annual average	Annual Difference	Percent Change and 95% CI (%)		Linear test p-value	Permutation test p-value
Traffic Stop Profile								
Total Safety Stops	13,968 (100%)	34,930 (100%)	15,786 (100%)	+19,144	+121.3	(+17.1, +318.1)	0.0055	<0.0001
% Safety Stops	6,119 (43.8%)	23,786 (68.1%)	7,296 (46.2%)	+21.9%	+47.3	(+20.0, +80.9)	0.0001	<0.0001
% Regulatory & Equip. Stops	6,073 (43.5%)	9,583 (27.4%)	6,951 (44%)	-16.6%	-37.7	(-54.6, -14.5)	0.0012	<0.0001
% Discretionary	1,776 (12.7%)	1,562 (4.5%)	1,367 (8.7%)	-4.2%	-48.4	(-55.5, -40.1)	<0.0001	<0.0001
Measures of Traffic Stop Disparity								
% Black non-Hispanic Stops	56.8%	54.7%	58.8%	-4.1%	-7.0	(-8.9, -5.0)	<0.0001	0.250
Black non-Hispanic TSRR	2.5	2.2	2.8	n/a	-21.3	(-28.5, -13.3)	<0.0001	0.125
Motor Vehicle Crash Outcomes								
Crashes (all)	5,298 (100%)	5,160 (100%)	5,925 (100%)	-765.0	-12.9	(-37.5, +21.3)	0.4439	0.125
Crashes (w/ injuries)	1,886 (35.6%)	1,639 (31.8%)	2,118 (41%)	-479.3	-22.6	(-48.5, +16.3)	0.2763	0.125
Traffic Fatalities	62.3	48.8	68.0	-19.3	-28.3	(-64.1, +43.2)	0.4146	0.125
Crime Outcomes								
Violent Crimes	1,223.6	1,233.5	1,257.3	-23.8	-1.9	(-32.8, +43.2)	0.9218	>0.99
Violent Crime Rate (per 1,000)	730.5	596.9	582.4	+14.5	+2.5	(-14.0, +22.2)	0.7815	0.750
Index Crimes	13,367.4	11,658.0	12,896.4	-1,238.4	-9.6	(-24.5, +8.2)	0.2923	0.500
Index Crime Rate (per 1,000)	7,848.1	5,637.3	5,933.4	-296.1	-5.0	(-12.8, +3.5)	0.2482	0.750

Table 7.4 Treatment vs. synthetic control: stop profile, crash outcome, and crime outcomes.

Table includes both annual averages pre-intervention (2002-2012) and post-intervention (2013-2016). Note: confidence intervals are not symmetrical around point estimates because different methods were used to produce each and small numbers further limited convergence.

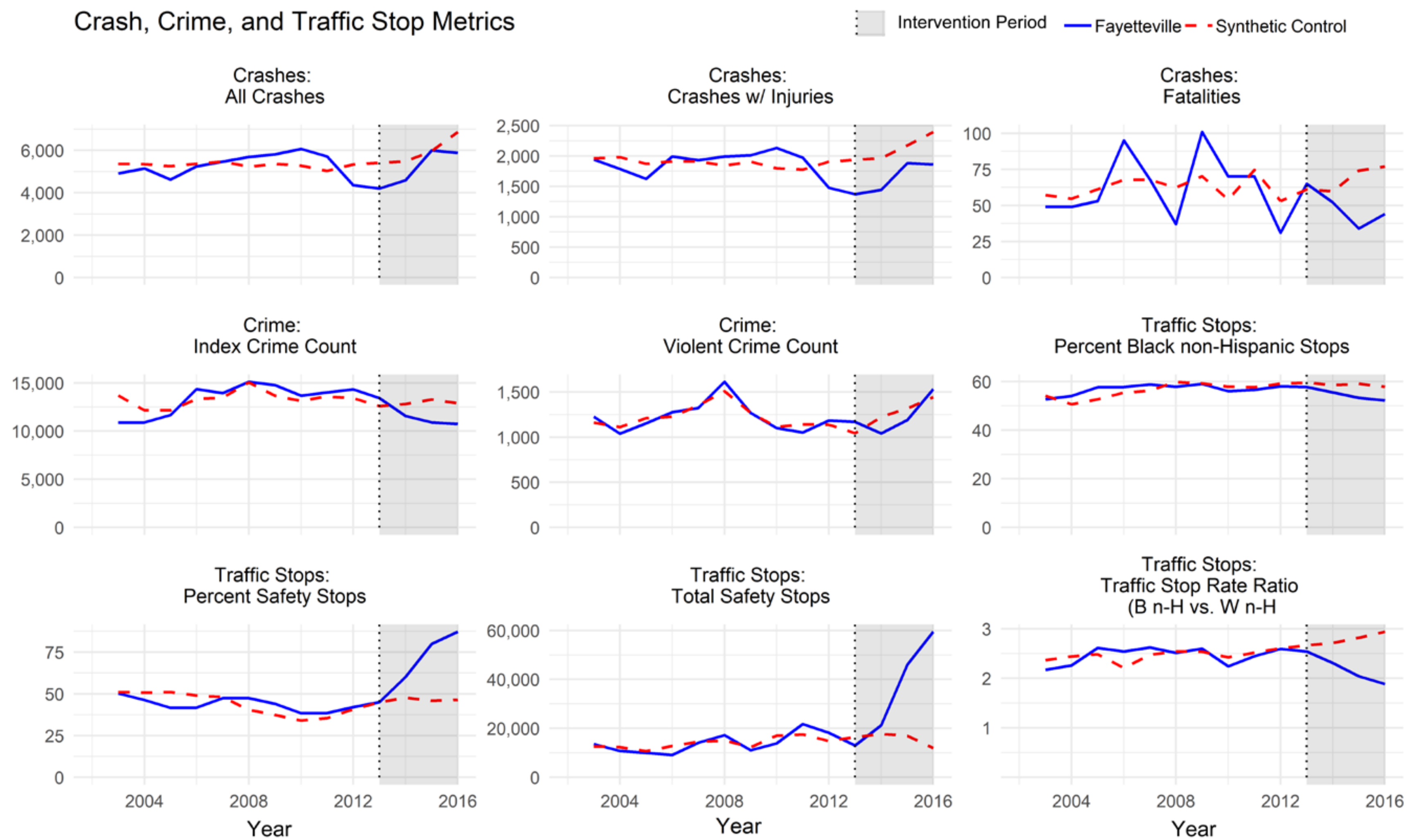


Figure 7.1 Crash, Crime, and Traffic Stop Metrics pre- and post-intervention period.

Fayetteville Police Department is compared to a synthetic control department built by the 8 most similarly urban, high population, North Carolina police departments best matched for the pre-intervention period.

7.5 **Discussion**

Traffic stop profile measures confirmed the implementation of the intervention strategy. Both the relative percent of safety stops and the absolute number of safety stops completed marked increased in Fayetteville in comparison to the measure-specific synthetic control agencies. This increase in the percent of safety stops was matched with a corresponding relative reduction in economic and investigatory stops.

Motor vehicle crash outcomes were all reduced, though confidence intervals were relatively wider. Measures of traffic stop disparities were also reduced, suggesting a focus on safety stops (and relative de-prioritization of investigatory and economic stops) was a viable strategy to reduce Black non-Hispanic disparities in their traffic stop program.

Neither index crimes nor violent crimes changed appreciably during the intervention relative to the synthetic control agencies: three measure point estimates saw small reductions and one saw a small increase, but these nominal changes were much smaller than their associated confidence intervals. This study does not provide any evidence of a negative effect on crime for de-prioritizing investigatory and economic stops. However, a more detailed view of the trend of the reduction in the total number of stops during the transition into the intervention suggests the first half of the Ferguson Effect, a reduction in output by some officers in response to community outcry and public attention, may have occurred. Staffing changes as agency culture changed may also have occurred during the intervention roll-out period and produced or contributed to this reduction in output as well.

These results suggest redesigning a traffic stop program for public health impact may reduce negative motor vehicle crash outcomes, simultaneously reduce some negative consequences of traffic stop programs (e.g. race-ethnic disparities, reduced economic stop burden on communities), and the relative de-prioritization may not have an significant impact on crime rates. Safety traffic stops, especially when directed at high crash areas using regular review and traffic stop GPS data for evaluation, may be a more effective public safety tool than economic or investigatory stops. If investigatory stops can be de-prioritized with little impact on crime, but carry with them negative consequences to community trust, those traffic programs may be de-emphasized even without a relative prioritization of safety stops.

However, these apparent public health wins can be fleeting, as transitions in administrators may bring entirely new or adjusted priorities. Since Chief Medlock's retirement in 2016, the percent of safety-related stops has dropped and the percent of Black drivers stopped has increased [104]. Future analyses may explore whether these new changes are associated with increases, decreases, or neither in crash, injury, and crime measures. Adherence to consistent public health priorities, especially when those relative priorities and implicit logics are made explicit, may help administrators transition while keeping interventions consistent.

Negative consequences of traffic stops

This study posits a relationship between certain stop types and public health outcomes under a conventional framework. However, that conventional framework ignores or downplays the real, negative consequences of traffic stop enforcement in practice. Regulatory and equipment stops, and their associated fines, are a direct form of criminalizing individual and community economic poverty. Beyond the immediate impacts, the harm of economic stops creates a negative spiral operating within communities collectively and individuals specifically, extracting wealth and people's bodies from low-income communities as the inability to pay mounting traffic tickets escalate into denied registration and warrants for arrest. The United State

Justice Department Civil Rights Division cited this extreme and racialized extraction of wealth through traffic stops in its review of the Ferguson Police Department [132]. When used unaccountably (e.g. without recording GPS data, as is the norm in NC), moving and safety violation stops can be enforced in an area with few motor vehicle crashes to justify them. Lastly, investigatory stops may have strikingly low contraband hit rates or racialized application [16], which subject some to antagonistic law enforcement interactions over years [106] without contraband to show for the interaction.

Beyond the serious financial and carceral consequences, at their most severe, traffic stops can have fatal consequences for motorists, even when unarmed. Sandra Bland, an unarmed Black woman who died in jail after a routine traffic stop, had multiple other unpaid traffic tickets at the time of her arrest, including for operating a vehicle without a license and lack of insurance [82]. Walter Scott, an unarmed Black man, was shot to death, in the back, by a South Carolina police officer after a traffic stop for a non-functioning brake light [8]. Philando Castile was pulled over forty times, for reasons including speeding, driving without a muffler and not wearing a seat belt, in the years running up to his fatal shooting during a traffic stop [106]. An uncritical increase in traffic stop enforcement means increased interactions with law enforcement, creating more opportunities for escalated and fatal encounters that may disproportionately impact low-income people and people of color given structural disparities and implicit bias. The associated loss of community trust has real public health consequences, including fewer calls for timely emergency services [34]. Beyond the negative consequences acknowledged to be more objective, public safety interventions driven by traffic stops should acknowledge the disparate, subjective, emotional experience drivers of color experience. Recent studies now document how these disparities in chronic stress get biologically embed (i.e. “get under the skin”) and have measurable and negative consequences for individual health [68,86,102], including specifically symptoms of post-traumatic stress order associated with increased interactions with police [70].

Program effectiveness, program efficiency

Central to this discussion are questions of absolute and relative intervention efficacy and efficiency. In Fayetteville's case, their safety stop program was likely more efficient because of its use of crash data to inform prioritization of intersections and the geocoded stop data to ensure intervention fidelity. However, safety related traffic stops are not the only method to reduce motor vehicle crash injuries. The efficacy of even maximally efficient traffic stop programs must be weighed against strategies from other sectors such as public education campaigns and built environment investments, which may be either or both more efficacious and cost-efficient [96]. Likewise, focusing on policing interventions for public safety in the absence of infrastructure improvements, given historical (e.g. redlining) and present disparities in those investments raise equity concerns [118].

When considering equitable investment in communities, this intervention to reprioritize traffic stops may best be a stop gap response to immediately reduce disparities and promote traffic crash outcomes but is not an ultimate solution. Though the intervention reduced racial disparities in Fayetteville compared by 21% of what they could have been, Black drivers still experienced over twice the incidence of traffic stops per vehicle miles traveled as White non-Hispanic drivers at the end of the study period. If not considering alternative interventions that may be more efficient, efficacious, or equitable, an investment in traffic stop programs in isolation may be capable of reducing motor vehicle crashes further but may require a totalitarian police state model stopping nearly all drivers for every possible infraction. Intervention considerations should include not only comparison of the positive efficacy and financial cost of programs but should weigh the negative collateral or intentional damages done. Instead, traffic stop programs may be intentionally phased out or scaled back alongside infrastructure investments and other interventions that carry fewer negative and inequitable consequences to remain in alignment with public safety needs.

The same principles are true when considering other public safety outcomes: though policing has seen large funding increases and expanding scope of practice [69], policing should not be seen as either a panacea overall or the most efficacious intervention for non-vehicular crime and injury specifically. Police do not replace mental health workers, social workers, or public health workers capable of implementing evidence-based programs at the individual and community level for substance overdose and violence-related outcomes. As law enforcement agencies are increasingly accountable to the efficacies and efficiencies of their programs, it is in their best interest to focus on programs, including traffic stop programs, that have fewer negative consequences, more equitable outcomes, improved efficacy, and efficient implementation when compared to interventions from other sectors.

Program priorities and the relative worth of life

In both law enforcement and public health, we implicitly and explicitly prioritize certain causes of disease, injury, and death over causes. Our prioritizations are reveals by our evidence and assumptions of efficacy and efficiency, program funding and implementation, and ultimately community investments enabled by political power. Even ignoring other sectors and intervention strategies besides traffic stops, police may compare the cost and efficacy of traffic stop programs in preventing injury and death by motor vehicle crash to preventing injury or death during a burglary, assault or homicide. When considering who is targeted by interventions, public health recommends considering distributions across population subgroups of both the burden of traffic stop preventable injuries and the exposure to traffic stops in the form of patrols patterns and priorities [138] along with efficacy and cost. Because of unequal distribution of outcomes, exposure to interventions, differences in intervention effectiveness and efficiency, these priorities come to represent the relative value of lives by race-ethnicity and socio-economic position. As example, if community investment (including through law enforcement and traffic stop patrol programs) in preventing deaths by assault grossly outweighs investment in prevention of deaths by motor vehicle crashes, overdose, or heart disease, and especially when the underlying burden of assault injuries and mortality is comparably low, we implicitly priorities the health and lives of populations seeking to prevent assault over other public health priorities and other populations.

These prioritization dynamics operate at multiple levels above and within agencies: within agencies as individual officer, patrol team, and precincts patterns; and above as clusters of agencies, statewide, nationwide, and between countries. At the national level we see these prioritizations in the focus on criminalizing drug use and addiction in urban, Black communities in the 1980s that lead to disproportionate incarceration of Black people at a level rarely seen anywhere else in the world [69]. In contrast, the multiple phases of the opioid epidemic since 2000, hitting more (but not exclusively) rural and white communities, has been comparably treated as a public health crisis rather than a criminal justice one [13,79]. Though this intervention analysis provided some contextual factors at the agency level, future research should not be limited to either implicit bias at the individual or policy effects at the agency level, but instead should continue to focus on questions or program priorities and implicit worth of human life at multiple and interacting levels.

Whether legally defensible or not, traffic stop programs may still be considered unjust and burdensome and may ignore racial disparities in financial hardships, eroded community trust, embodied community stress, and injury and loss of life outcomes in some communities to promote or appear to promote the well-being of other communities. Even within the same community, for example, a seatbelt program that extracts large amounts of financial resources may cause serious harm to individual and community health and may outweigh the injury prevention benefit. Co-designing traffic stop programs along with impacted communities may alleviate some, though likely not all, of these negative outcomes, given there are multiple underlying dynamics at play [122]. It is precisely these implicit disparities in the value of people's experiences, bodies, and ultimately that drives associated policy platforms calling for the end of criminalization and dehumanization of Black and low-income communities [6].

Accountability

We argue that public health has a fundamental interest in detailed traffic stop data given associated public safety outcomes under the conventional frame and equity considerations under anti-racist frameworks [47]. However, not all states maintain active traffic stop databases like North Carolina's. Further, most active traffic stop databases that do exist were started recently. When contrasted with many other public health surveillance systems, these limited data suggest a relatively limited oversight of law enforcement activities and adverse events in some communities. Public health has acknowledged that data on deaths by officers, specifically, are public health data, can and should be maintained [44,86], and that collecting law enforcement data in general fundamental to accountability and trust [88]. Data collection on traffic stops should also include some within-agency spatial component, as Fayetteville has elected to collect, such as spatial coordinates or an address or intersection that could be retroactively geocoded. Such detailed data on traffic stop programs also benefits police agencies. Spatially-referenced traffic stop data can inform prediction and intervention models of public safety events like crashes and violent assaults, and property crime, and also ensure accountability within the agency and to community priorities. For instance, the National Institute of Justice and the Bureau of Justice Assistance collaborated with the National Household Traffic Safety Administration to promote Data-Driven Approaches to Crime and Traffic Safety (DDACTS) [31]. DDACTS includes a series of workshops, an associated journal, and techniques to formalize hot spot analysis of incidents and crashes. GPS tools are increasingly low-cost, included in most cell phones, and retrospective geocoding are inexpensive.

As example of public safety interventions and equity implications, the National Household Traffic Safety Administration (NHTSA) put out a manual for state highway safety offices, outlined evidence of law enforcement activities including types of traffic stop [57]. This document drove updating of CDC guidelines around motor vehicle safety interventions [72]. Included as an evidence-based intervention are "a saturation patrol (also called a blanket patrol, 'wolf pack,' or dedicated DWI patrol)" [57]. Likewise, movement from secondary to primary enforcement of seatbelt laws (e.g. allowing seatbelt ticketing when no other infraction is present) is associated with more seatbelt usage and reduced traffic crash fatalities. But when public health advocates for saturation approaches do not acknowledge, the new approaches may disproportionately burden under resourced communities with the negative consequences of traffic stops. And, without some within-jurisdiction accountability, agencies are free to use their discretion to distribute DWI and seatbelt patrols into neighborhoods that may not have the political and economic capital to fight in court and may not equitably weather the negative effects of such saturation.

Limitations

This study has multiple limitations. Since only one agency enacted the intervention, our findings are suggestive but limited by sample size. If a group of agencies were to adopt this prioritization results may be more robust. We hypothesize that the synthetic control method improved confounding control compared to a difference-in-difference model. However, an approach that incorporated data on more agencies and more covariates under a more detailed confounding control scheme would likely produce more accurate results than our approach of matching on the pre-intervention period. That said, particularly when there is a scarcity of implementation sites and promising interventions, documentation of aspiring anti-racist interventions is worthwhile in the face of these limitations [77].

Further, the capture of race-ethnicity in administrative datasets has known limitations [83]. Race-ethnicity is a powerful social construct associated with many associated health disparities [130], so many we that require dedicated frameworks to harmonize them [37]. Because of its social construction[46], the meaning of race-ethnicity changes over place and time and can vary person to person even within the same time and place. Health research acknowledges that self-identification may differ from social-identification [76]. Even in the same person, conceptions of race-ethnicity change over the life course [93]. Concretely in this study, the self-identification options in justice databases are limited and may not match driver's self-identity. Stopping officers may not refer to driver-specified race-ethnicity (notably incomplete in NC driver's license records [111], but instead fill out form SBI-122 based on their own ascription of the race of the driver. Indeed, there is documentation that in some regions law enforcement officers may knowingly misidentify race-ethnicity in response to scrutiny under new racial profiling laws and accountability that databases would seek to provide [127].

Conclusion

Reprioritizing traffic stops for public health can reduce negative crash outcomes, reduce disparities, and may not have negative impacts on crime. More generally, a public health anti-racist approach requires, for example and at least, that injury prevention researchers who design interventions that will be enacted by law enforcement (e.g., seatbelt traffic stop campaigns) to consider the reality that some agencies and officers may intentionally or unintentionally target populations in racially disparate ways. The collateral damage of even well-intentioned public safety interventions may outweigh their benefits. These damages may be disparately born by low-income and communities of color. Public safety and public health are intimately related endeavors, as evidenced by this demonstration of their relationship around traffic stops. When engaged with public safety issues, public health should adopt a critical view of policing at the same time both fields must critically interrogate their own historical and present-day practices. Conventional logics, such as the Ferguson Effect belief that de-prioritizing investigatory stops is associated with increases in violent crime, may not hold up to critical scrutiny.

Public health has outlined an explicit call to anti-racist practice and principles. Law enforcement organizations, individual law enforcement agencies and officers, city councils, county boards, and community groups may elect to take up that call to guide their own activities. When co-designing traffic stop programs, these groups should consider goals of equity and maximizing public health impact alongside effects on community trust. But regardless of law enforcement agency action or non-action, public health advocates can use traffic stop datasets to both ensure their efficacy for public safety goals and document and act on any racially disparate impacts of these programs.

CHAPTER 8 - DISCUSSION

8.1 **Study Strengths**

The NC traffic stop and search dataset is one of the most complete in the nation (Baumgartner et al., 2019), making it an ideal setting for examining nuanced stop rate and population questions. Few states require centralized reporting, and of those, many allow reporting in different formats using agency-specific forms. The North Carolina State Bureau of Investigations requires agencies policing jurisdictions of a minimum population threshold [23] to report stop data on a single state form (SBI-122) that has been consistent since 2002, providing over a decade of consistently formatted data. Though the number of police agencies fluctuates slightly over time, this dataset includes 308 of the 518 city, county, state and place-based (e.g. school, hospital, etc.) police agencies in the state. Though representing only 60% of the agencies, these 308 agency jurisdictions with data represent an underlying residential population nearly 100% of the state population (Figure 3, right).

The cross-disciplinary UNC-CH team is uniquely positioned to lead this research. The team includes epidemiologists and public policy experts who have been working on these questions in this dataset for years. Further, because of the immediacy and demand for this topic, the research team has developed existing relationships with community groups like the NAACP, advocacy organizations like the Southern Coalition for Social Justice, and numerous police agencies, both local and statewide. Preliminary analysis in this dataset focusing on searches has gained significant traction in both local and national press [12], suggesting the time is right and this research is overdue. These policy and policing relationships are essential for understanding the practical realities of the dataset, pressing questions of community members and police agencies, and the policy options being considered. The research team is continuing to innovate with related questions, working with specific police agencies who have volunteered more data to pilot new hypotheses and increase the geospatial precision of the data to investigate related public health questions. The research questions in this project will be tested and improved by these relationships and opportunities.

More specifically, recent literature addressing the specific gaps in accurately estimating race-based police traffic stop rates has advocated for specific methods that will be used in this project, including bringing in novel, supplemental data sources like not-at-fault drivers and vehicle registrations. These methods, combined with a more than large stop dataset, have a high likelihood of successfully coming together to produce meaningful estimates and concrete best-practices.

8.2 **Study Limitations**

Though there are many strengths to this analysis, there are outstanding limitations. Here I list four: (1) the administrative data used to inform both aims, (2) the lack of a gold standard for rate building used in both aims, (3) the need for generalization and the tyranny of power and small numbers, especially with sub group analysis, and (4) theoretical limitations known, but unanswered by this research.

SBI-122

The study outcome is being stopped by police within a jurisdiction. This outcome is assessed through police submission of a consistent form (SBI-122). There are known limitations of this outcome assessment tool. However, it is not only the only outcome assessment method available, but also currently the best state-wide dataset in the nation for assessing this outcome. These limitations include the following: Though a few jurisdictions use point-geocode data, the current form (1) does not record the location of stop (other than within a specific agency's jurisdiction), (2) does not record the resident location of the driver (only demographic information), (3) is universally used but not universally recorded (the smallest city agencies are not required to report their data), (4) is used for nearly all stops, but not all stops within an agency (NC general statute exempts G.S. 20-16.3A roadblock and checkpoint stops), (5) inter-agency differences in coding (across tens of thousands of officers) create data errors.

However, this research was still be accomplished within these limitations. The missing data from small agencies and select stop-types not included was very small (<1%), and we hypothesized it to skew findings of racial disparity in a known direction (away from the null). This research team is working in collaboration with police departments, and has been for over two years, which allows some feedback around data-limitations. The key variables needed for the outcome assessment at the agency-level of analysis were robust enough to these limitations to permit analysis.

Generalization and the lack of a gold standard

I am using publicly available versions of multiple datasets that have smaller-area data available. NHTS has block-identifiable data available and the census ACS commuting estimates could be requested at the micro-level. I have done this both to simplify the analysis and to emphasize techniques more available and realistic for generalizable use. However, better models would account for this limitation by using distributions for my estimates across models taken from the aggregate datasets instead of rederived more locally from the record-level data.

This project is limited in scope. While building a national model (see Areas for Future Research, following) is important, this method of improved rate estimation is not entirely replicable to other states. However, the demonstration of the degree of difference when considering travel disparities should serve as motivation to continue this research on how to generalize these methods for small-area estimation of driving by race-ethnicity and other stratifications.

Moreover, in the absence of a gold standard, it was difficult to conclude the sensitivity analysis in Aim 1 with a clear sense of which method was more valid. Larger national studies that generate VMT estimates by wholly different methods may provide some increased certainty that models are accurately estimating the same driving realities, but without a gold standard, different models may also be incorrect in the same way.

Group stratification and small numbers issues

Beyond generalization, issues of power and small numbers arise at multiple levels. First, both Aim 1 and 2 limit their analyses to certain select race-ethnicity groups. Though Aim 1 produces adjustment factors for Native American / Indigenous populations, because of small numbers across multiple agencies it does not provide rate comparisons between models, even though this population experiences disparities in many places similar to Black non-Hispanic and Hispanic groups explored. Of note, small numbers in this case is partly due to historical genocide. A strict adherence to the primacy of study power systematically excludes the most marginalized and small populations. This creates equity issues as we seek to stratify to more accurately describe group adherence.

Relatedly, grouping many unique populations into the "Asian" group in Aim 1 likely group together disparate experiences: a hypothetical family of recent immigrants from Burma who do not speak English and live in subsidized housing in a low-income areas may have very different experiences of law enforcement as than a compared to a Chinese-American family who immigrated two generations prior and have family members with graduate degrees. While none of these hypotheticals are effective stereotypes or accurately summarize an average experience of a particular racial-ethnic group from Asia, it is important to note the size and diversity of populations are erased when lumping disparate groups.

However, focusing on the social construction of race, it may be that in some areas, while Chinese, Japanese, and Korean communities and cultures may not intersect, just as Portuguese-speaking Brazilians and Spanish speaking Mexicans may not intersect, their treatment on an individual level by officers and at the neighborhood level by agency patrol patterns may follow similar White implicit biases and agency decision making. Meaningful race-ethnicity distinctions to individuals and groups, internally, may not carry to social experiences, externally. Mathematically, larger groupings may enable models to provide more stable estimates. However, those average estimates may not apply to any subgroup if the subgroups measures (e.g. rate ratios) are widely distributed around the group average. In that case, average estimates may be more misleading than producing none at all. These are challenging nuances to draw apart in research.

Defining the Fayetteville intervention

There is more nuance to the Fayetteville intervention than can be conveyed by any one, or even multiple, quantitative measures. The percent of safety stops, the priority metric used in the Chapter 8 manuscript, is an indicator measure that demonstrates a significant change in policy and practice. However, this should not be confused with it being the sole mechanism of the intervention and it does not capture qualitative dynamics that made the intervention challenging to implement. Instead, an agency likely needs to do more to follow Fayetteville than simply increase the percent of traffic stops to enact this intervention.

The figure below demonstrates three related measures traffic stop measures that capture quantitative facets of the intervention implementation: the percent of safety stops, the total number of safety stops, and the total number of stops. As the intervention began, 2013 saw a marked drop in the total number of stops just before and in the first year of the intervention. This corresponded with marked critical attention on the activities of the police department and large staffing changes of not only the head administrators, but some of the officers as well. And these priorities were dramatic shifts quantitatively: not only did the percent of safety stops change from a low of 30% to over 80%, the raw number of safety stops increased by a factor of three by the end of the intervention period. Interviews with administrators the drop in the number of stops may have partly been due to a negative response in some officers to the increased accountability, negative attention, and perceived challenges to community trust. Qualitatively, interventions may require organizational and cultural change work that was required to shift priorities so dramatically; moreover, the safety stops themselves, thanks to the GPS data collected, may have been more effective. Quantitatively, this may be framed as an intervention ramp up period. None of these dynamics are captured by the single percent safety stops measure. Percent of safety stops is therefore a representative indicator of the change, but again, as a single measure does not fully capture the full realities of this implementation.

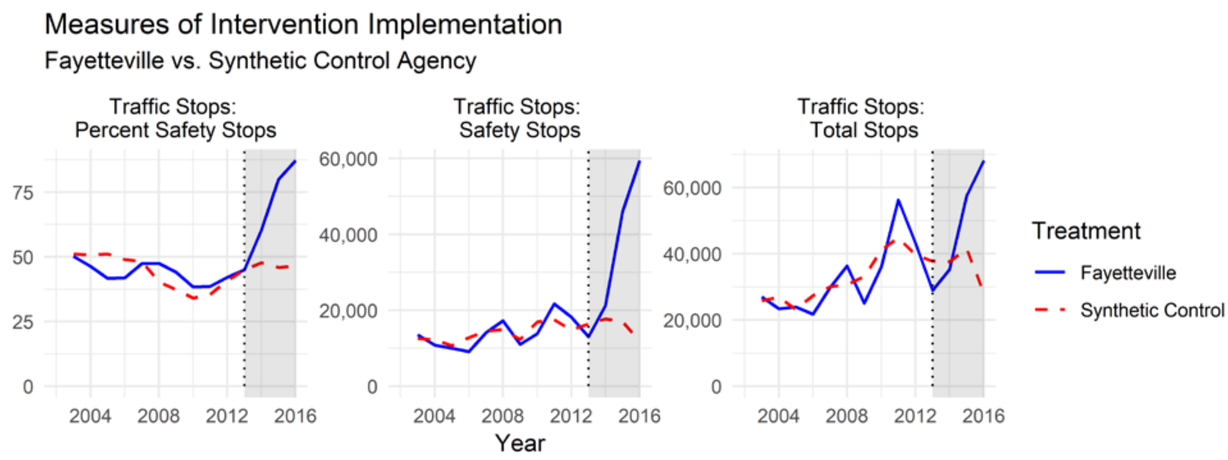


Figure 8.1 Measures of intervention implementation, Fayetteville vs. synthetic control.

Theoretical limitations

Lastly, given the multiple theoretical frameworks mentioned, it will be difficult to do right by them all. This project attempts to be both an academic dissertation and community project, to be harm reductionist while not abandoning deeper alternatives to law enforcement as it stands. This will be a project of compromise, but it is my hope that it's stronger for it rather than weaker.

As discussed previously, traffic stops are a fundamentally multi-part and multi-level phenomenon, though this dissertation's analyses focus on the agency as the unit of analysis and action. Alternately, previous studies have taken alternate vantage points within that multi-level framework; some, for instance, focus on training individual officers as an important component of policies to address disparities (Banakou et al 2016; Assari 2018; Saywer and Gampa, 2018; NYC 2018), and some tests of disparity (RTI STAR) have focused on this individual and interpersonal level. To combat a tendency to focus on individual, behavioral-level conventional frameworks, this study intentionally situates itself at the individual level. But other studies could place such an agency level analysis within regional or statewide models, and then further account for individual level analysis, to better capture these multi-level dynamics. Studies like that could likewise better capture neighbor-level dynamics, like spatial proximity or lag effects, that may have explanatory power.

Lastly, this study has acknowledged limitations, as explored in Chapter 6, when compared to anti-racist study principles described in the Public Health Critical Race Praxis (PHCRP). While some of the PHCRP framework is maintained as a thread throughout both aims (race consciousness, primacy of racialization, race as a social construct, non-biological construction of race-ethnicity and the racialization of meaning, structural determinism), other praxis values are largely underexplored or wholly ignored (gender as a social construct, intersectionality, disciplinary self-critique, voice). Future studies could do better.

Racist extraction of wealth from black communities, bodies.

8.3 **Anti-racist self-critique**

The same frameworks used in Chapter 6 to contrast conventional frameworks from critical frameworks, used throughout the dissertation, can be used as a self-critique tool for the dissertation itself. Structured self-critique enables a systematic way of identifying limitations, acknowledging imperfections, and learning from mistakes. Below are self-critiques for this dissertation organized within two frameworks: (1) the 11 principles of PHCRP and (2) the 15 characteristics of White Supremacy Culture. The PHCRP definitions were reviewed in detail in Chapter 6. For convenience of readers, since these definitions were not reviewed in detail, a link to the characteristics of White Supremacy Culture given here in a footnote [3]. I offer an admittedly rough self-score, though quantifying these aspects in a rubric may have limited use for others. This self-critique is not comprehensive: subsequent reflection with dissertation committee, teachers, community collaborators, and peers will undoubtedly reveal other opportunity for reflection, learning, improvement.

PHCRP Principle	**Self-Critique**	**Self-Score**
1. Race consciousness	The topic was chosen and analysis designed by author's consciousness of own race privilege and the dynamics of structural racism after years of collaboration around issues of race disparities in criminal justice outcomes. Race is primary to analysis, discussion, and author’s position, not an afterthought.	4
2. Primacy of racialization	Dissertation centers disparities, and discusses primacy of racialization, but does not center as much. Aim 2 provides little discussion of cultural pushback in law enforcement agencies for attempts to reduce racial disparities.	3
3. Race as a social construct	Dissertation acknowledges social construction of race in discussion, especially the meaning of the race-ethnicity variables in administrative data and the malleable role of identity across space and time. Little treatment of race and traffic stops historically, though relevant material exists (e.g. history of race construct through important legal cases, racialized history of traffic stop programs.	3
4. Gender as a social construct	Only brief references to interaction of gender and race ethnicity, though literature (on both traffic stop disparities and underlying driving patterns) provides rationale for a deeper dive.	1
5. Ordinariness of racism	Traffic stops and their disparities are described as pervasive and consistently high. Any possibility of neutral traffic stop programs is critiqued. Intervention (Aim 2) is offered as a possible harm reduction technique, with some critical discussion of the limits of incremental redesign as a strategy to reduce racism's effects.	3

[3] http://www.dismantlingracism.org/uploads/4/3/5/7/43579015/whitesupcul13.pdf

6. Structural determinism	Behavioral focus is critiqued, and traffic stop programs are described in the context of structural racism and disparities across criminal justice and public health. Little attention is given to other structural factors that may explain or co-occur with traffic stop disparities.	3
7. Social construction of knowledge	Preface and discussions acknowledge the collective (vs. individual author) construction of aims and underlying concept relations. Community coalitions helped to shape conceptual framework. Some attention given to alternate dissemination techniques (public website, public forums). Little attention given to how communities might create alternate narratives with explanatory power besides those posited in manuscript 2.	3
8. Critical approaches	Significant time spend in discussions and analysis choices to critically examine racial disparities in underlying factors beyond conventional frameworks. Nominal critical treatment of limits of incremental intervention design and deeper redesign / abolishing of policing as we know it.	3
9. Intersectionality	Brief mention of gender x race intersection in discussion and lit review. Some serious treatment of interrelatedness of income and race disparities in both Aim 1 and 2.	3
10. Disciplinary self-critique	Some acknowledgement of public health interventions context of colorblind interventions (e.g. seatbelt enforcement, police collaborations) that have little consideration for racial equity. Self-critique of dissertation using this tool.	2
11. Voice	Ongoing collaboration with Black lawyers, experts and activists helped shape dissertation. However, bulk of design choices were made before completion of even baseline examination of problem using PHCRP / CRT, so theory is mainly an afterthought for evaluation. Some, but minimal effort was made to, for instance, hear local stories about the Fayetteville intervention by resident experts of color. Specific aims were designed to prioritize research that might be used for action based on a wider, interracial accountability group. However, in order to fit it into a dissertation, some priorities may have been compromised. Though dissertation was intentionally informed by frameworks from authors and organizations of color, dissertation still largely feels like a White (albeit aspiring anti-racist) voice using White methods, even if critical of some white logic.	3

Table 8.1 A Public Health Critical Race Praxis self-evaluation of this dissertation.

Scoring: 1 = little to no discussion; 2 = acknowledgement, but little treatment; 3 = discussion and some analysis considerations; 4 = study design & analysis choices driven by principle; 5 = thorough and complete application.

White Supremacy Culture

Characteristic	Self-Critique	Self-Score
1. Perfectionism	While I felt constantly unnerved by this dissertation's imperfections, the quality of my work, and timelines, I also acknowledged (with help) that while quality is important, this dissertation is not my magnum opus. Even if it was, it would be imperfect.	3
2. Sense of Urgency	I struggled with a sense of urgency through this project. I found it regularly difficult to balance "why hasn't this work already been done?" and "this work is needed… yesterday!" with the timelines of dissertations, competition projects and priorities, and the classroom component of my PhD and second masters. I regularly am challenged by this characteristic, even as I am paradoxically fed by a sense of energy and immediacy in my work. Particularly after dates were set, I procrastinated to some degree, feeding on the sense of urgency as those dates got closer.	2
3. Defensiveness	I have actively worked to reduce my defensiveness not only in receiving committee feedback, but also peer and community collaborator feedback. I feel I have not defended Whiteness in general. I have found myself confronting internalized feelings of defensiveness given this dissertation has not been top priority in my life, but one juggled against family, community, and personal obligations.	4
4. Quantity over quality	This dissertation is long, most likely too long. I have approached the problem from many directions using many techniques, and still not exhausted what feels like appropriate breadth. For as many supplemental analyses as are included in this analysis, an equal number were cut. The quality of the entire dissertation would have been better with fewer analyses and less, but more honed, writing. I believe I have erred on the side of breadth over depth, quantity over quality here - though this is a difficult characteristic to balance, paradoxically, with perfectionism.	2

5. Worship of the written word	A dissertation may be one of the most essential examples of worship of the written word. I do not believe the dissertation, as a piece of writing, has fundamental power to change. I have acted on that belief by collaborating with community groups, discussions with law enforcement administrators, building fact sheets and deliverables (still written), and supported other dissemination tools like websites. Organizing is fundamental to change, and relationships (not just written artifacts) are fundamental to organizing. I believe I have balanced that well given the constraints of an academic dissertation.	4
6. Only one right way	Pursued many methods and possible aims before settling on this way, which does not necessarily feel like the right or only way. Humbly unclear whether dissertation Aims as they are will practically bring benefit, even if they're designed to make that possible. Other strategies for dissemination and change tried alongside peer-reviewed scholarship.	3
7. Paternalism	Actively avoided prescribing behaviors or organizing strategies for communities, and acknowledged intervention is no panacea. Decision making on dissertation aims and design influenced by community collaborations, but ultimately were largely personal decisions, not community-led research.	3
8. Either/or thinking	Avoided p-value focused, either/or hypothesis testing frameworks, focusing on instead on magnitude of effects and continuous variance frameworks. Some resistance to "either include or drop" by deemphasizing and leaving superficial (without entirely dropping) prior supplemental analyses in Appendices.	4
9. Power hoarding	I have acknowledged ways in which the justice system hoards power when it acts without accountability by traffic stop rate and disparity dynamics, not electing to collect sufficient data, and not co-developing traffic stop programs with communities. However, beyond agency power hoarding, this PhD further concentrates power, in the form of both expertise and dominant culture accepted credentials, in one person (me). It is my responsibility to redistribute that power by enabling others to control it through my networks of accountability and service-partnership with community organizations. I have attempted to do this throughout this dissertation development and will continue to do this going forward but am increasingly at risk of power and privilege hoarding.	3

10. Fear of open conflict	Little open conflict in dissertation design and writing process, and disagreements were handled amicably and directly in conversation. Some open conflict in community meetings and with certain administrators. Lack of open conflict may part be positive conflict resolution, part avoiding challenging conversations.	3
11. Individualism	Larger dissertation-related processes saw author engaging as a team member with community coalitions as a partner. Dissertation benefitted greatly from formal feedback from committee and ongoing, informal feedback from peers and other teachers, and connection with dataset maintainers. Dissertation still largely an individual endeavor, even if attempts were made to collectivize it.	3
12. I'm the only one	Little delegation - may be partly fundamental to conventional dissertations. Even in community work, jumped quickly to "what can I do?" instead of "what has been done?" Eventually found Frank's research group, and provided some collaboration, but did not effectively delegate (though was delegated to some small tasks).	2
13. Progress is bigger, more	My discussion acknowledges that a viable change strategy may be to shrink, not grow, policing, focusing on quality over quantity. On the writing front, though have been challenged by quantity-quality dynamics, I have benefitted from progressing the manuscript chapters, along with my dissertation committee, by cutting back hard and increasing quality, not growing those papers.	3
14. Objectivity	Wherever possible I have critiqued objectivity: my voice as an author, traffic stop program design, the nature of infractions and crime, the documentation and definition of race. I have tried to maintain a "Strong Objectivity" framework (Harding, 1995) that requires deep and person/institution specific context. I do appeal to a lay sense of objectivity in Aim 2 by linking traffic stop programs to "objective" measures like motor vehicle and assault injuries. In this case, I am using this "White logic" to critique itself, in keeping with frames from the Black Lives Matter movement that call into question whether lives are treated objectively of equal value.	4

15. Right to comfort	"Discomfort is at the root of all growth and learning" (Okun, 2000) certainly seems to be my experience here. Did not assume, nor act, as if this dissertation would be pleasant and without personally challenging moments. Did not add fuel to discomfort fire by being meta-dissatisfied with that discomfort. That said, for many of these principles, it's still important to not artificially produce discomfort by lax organizing or overreach; discomfort is not a proxy for meaningful work.	4

Table 8.2 A Characteristics of White Supremacy Culture self-evaluation of this dissertation.

Scoring: 1 = little to no application of antidotes, dominant internal and external experience of cultural component; 2 = acknowledgement, but little resistance; 3 = Mixed resistance and adherence to cultural characteristic; 4 = conscious, continuous internal resistance and application of antidotes; 5 = thorough and complete understanding and resistance.

To summarize these anti-racist self-critiques: while I and this dissertation have benefitted from prior training and the recent application of PHCRP, there is much room for improvement in both future research projects as well as my own internalized experience of White Supremacy Culture. With help from others I hope to continue to improve in these areas.

8.4 **Areas for Future Research**

This dissertation furthers the literature on traffic stop disparities and one possible intervention to reduce them. However, it leaves many questions unanswered. Here I list five specific next steps that this results suggests are useful research activities to better understand and act on traffic stop disparities.

8.4.1 National small-area estimation of race-ethnicity-specific driving denominators

The most pressing need is for a nationally scalable method to measure traffic stop rates. Aim 1 of this project demonstrates the need for travel informed denominators for traffic stop rates, especially when considering disparity measures. It calculates estimates of vehicle miles traveled for all agencies in North Carolina, then subsets to agencies with sufficiently stable rates and complete data to demonstrate that disparities may compound if vehicle access, vehicle miles traveled, and distribution of those vehicle miles share the same disparities.

However, this method does not scale without modification to the United States. While NHTS is a national survey, only some states are weighted sufficiently to provide statewide estimates (Roth, Dai & Dematteis, 2017). Beyond that, individual agencies would be better off with sub-state adjustment instead of sharing statewide adjustment factors as was done in Aim 1.

Once a national model is built, it could be immediately disseminated. North Carolina has the nation's first open policing website (https://opendatapolicingnc.com), borne out of partnerships and data linkages from this UNC-CH research team. Community groups and police agencies have exploring its use, and other states are looking to this first iteration as a model. Lessons learned in describing police stop rates by race in this project, once generalized to a national model, can be disseminated not only through traditional channels, but also through future improvements to the publicly available website and other, future state websites.

However, such a project may require other datasets besides those used in this analysis for North Carolina. Those supplemental datasets might include some or all of the following:

- Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) block to block travel data

- License & registration data, statewide or nationally
- Census American Community Survey commuting and day-time population tables
- Bureau of Transportation Statistics (BTS) Local Area Transportation Characteristics for Households (LATCH) Survey

LEHD LODES

The Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) file describes the residence and employer location of. LEHD LODES is also produced by the census, representing a complementary data product to the Census ACS [58]. LEHD data derives from counts of jobs covered by unemployment insurance. In North Carolina, this represents the household-work travel habits of over 3 million North Carolina residents. These origin-destination start and end points can be used to model the through-travel pathways of those drivers, either by shortest Euclidean distance ("as the bee flies"; see figure, below) or shortest path along roadways (shortest "Manhattan" distance). This distribution could also be used as a guide to generate a fall-off buffer, and proportionally capture populations at risk and their demographics depending on the distance distribution of these origin-destination point sets for any given census block.

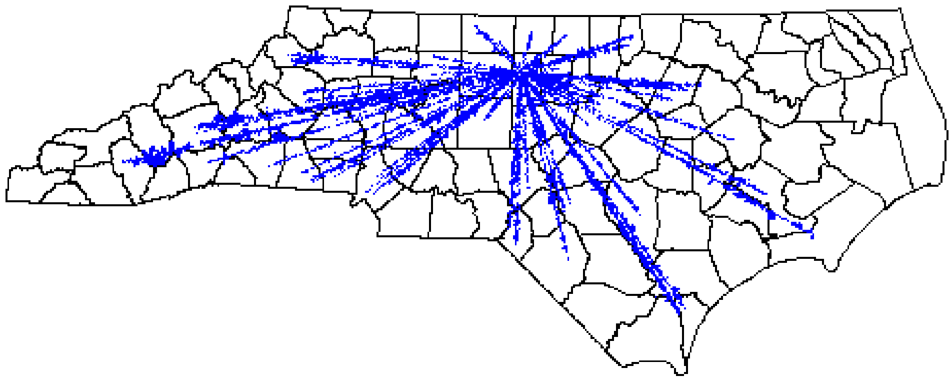


Figure 8.2 Simplified LODES demonstration: 250 origin-destination paths to a census block.

Origin-destination paths (assumed; not in LODES file) could also take shortest-distance along actual roadways and use a buffer to determine impacted roads. Analysis and visualization by author, 2015.

License & Registration Data

Though active licenses do not guarantee a potential driver has access to a vehicle, and some drive vehicles without licenses (both sources of bias in this dataset), license data may be a better proxy for driving than overall population (including children and those without vehicles or licenses).

Though I have contacted NC DOT and DMV directly, I have not made any headway in obtaining this aggregate dataset of valid licenses per LEA jurisdiction. The UNC Highway Safety Research Center had made separate efforts, but it is unclear when this dataset would be available in time for this research. Should it be available, it could be incorporated into this project as separate facet to the sensitivity analysis or in a combined model. However, due partly to a historical quick in North Carolina, race-ethnicity data is not required to be collected on drivers licenses, so this data is blank for many drivers [111]. Further, per Garrett & Crosier [53], an astonishing one in seven North Carolina drivers have active suspensions (over 1.2 million), even as their activities of daily living (caring for children, commuting to work, getting groceries, etc.) require vehicles. License data in states with large suspension programs may be uniquely ill-suited to estimate drivers.

ACS Commuting & Day-time Populations

Even adjusted for driver status and miles driven within a given jurisdiction, the resident population does not represent the driving population at rate of stop because of work commuting patterns. The American Community Survey (ACS) is a sample of around 3.5 million addresses annually that acts as a supplement to the decennial census [58]. The ACS estimates of commuting are based on a sample of workers 16 and over and provide estimates of the day-time population (in contrast to the residential, night-time population) of cities and counties. Preliminary data using census 2010 commuting / day-time population estimates suggests an obvious overall commuting pattern, where large populations drive from rural to urban areas. This change in population is significant enough to more than halve or double the population in many police jurisdictions (see figure, below). This change would not only clearly influence the magnitude of the individual rates, but if commuting patterns are even marginally different by race (which they likely are), it would be a source of significant influence on measures of disparity across race given the size of the overall influence.

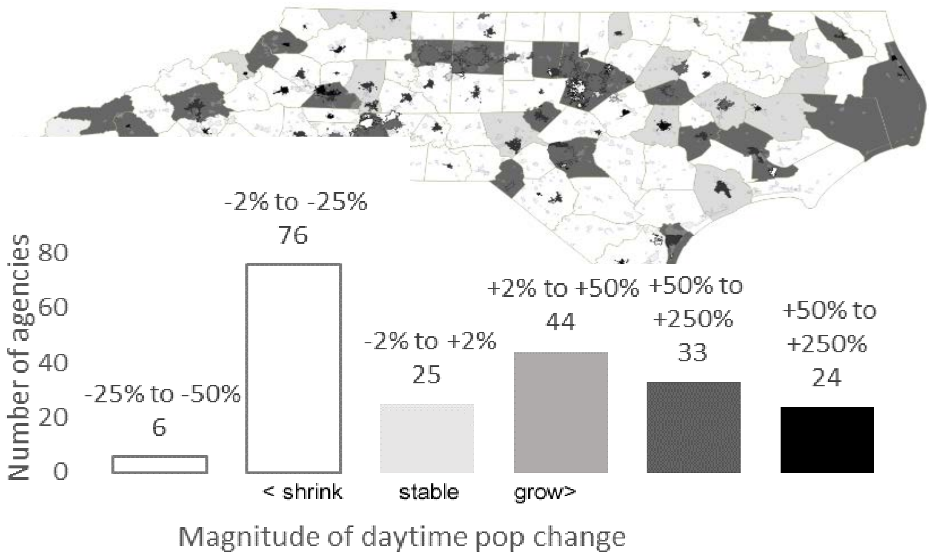


Figure 8.3 Changes in commuter (daytime) population-based agency populations.

Commuting patterns shift driving populations from rural areas to urban centers, halving (white) and doubling (Black) populations in many areas. These commuting patterns, likely disparate by race, dramatically change rates and risks of police stop in a given police jurisdiction. Analysis by author, 2015

ACS commuting estimates do not have demographic data, so cannot be directly used to estimate daytime populations. However, the total number of the daytime population could be used with place-specific buffers and the surrounding detailed census demographic data to estimate the demographics of that increased day-time population, producing a new, demographically adjusted at-risk profile. Commuting data is further limited by the fact that commuting is still a subset, however significant, of all travel. ACS day-time models therefore do not capture important in-vehicle activities like visiting family and friends, grocery shopping, travel for vacations, and social events.

BTS LATCH

The Bureau of Transportation Statistics (BTS) Local Area Transportation Characteristics for Households (LATCH) data may be the closest tool to providing estimates needed for a nationally consistent method of calculating traffic stop rates and their disparities. LATCH combines vehicle access and travel information from NHTS with census data at the tract level [22] to build a national model that incorporates urban/suburban/rural distinctions and US regional differences. That national model provides estimated weekday household person miles traveled, person trips, vehicle miles traveled, and vehicle trips for each census tract. This is a similar strategy to the small-area simulation completed in Aim 1, but the by modelling the NHTS survey results it can be extrapolated to small areas without using a single state-wide estimate for all small areas within a state.

While LATCH does not include race-ethnicity data, it may provide a useful basis a consistent, national method. Once built, tract level data on vehicle miles traveled by race-ethnicity (or other strata) could be aggregated into agencies similar to Aim 1's method of considering the primary patrol agency.

8.4.2 Interpretation of Stop Rate Variation Between LEAs

As explored in Aim 2, common interpretation of investigatory stop rate ratios by race suggest that they are driven by prior crime rates. If that is the case, prior crime rates should cause future stop rates, both overall and by demographic subset, perhaps with some reasonable lag and with appropriate confounding control. Therefore, jurisdiction-specific stop incident rates by race from Aim 1 could be used as ecological-level exposures for NC police agencies, with overall and race-specific incident crime rates as the outcomes, using Poisson-distributed general linear model regression for overall rates and multi-variate regression for race-specific rates. Covariates could include population size, jurisdiction type, or income distributions. Null relationships and outlier jurisdictions would be suggestive if crime rates explain little of the variability in investigatory stop rates and vice versa.

Similar exploratory analyses could be conducted using traffic-related injuries and race-specific income disparities. Each of these ecological questions respond to common police, community, and media interpretations of why specific agency stop rates, especially groups of stop types covered previously, are high compared to other agencies. Insignificant regression coefficients, correlation parameters, and outliers to those relationships contextualize common interpretations and suggest areas for future research.

Once the first aim is established, the second aim, exploring the meaning of these race-specific stop rates as markers of racial disparities in policing, could follow a traditional ecological, cross-sectional and time-series analysis using causal inference informed generalized linear model regression with a Poisson link for rate modeling. Geographic weighted regression (GWR) may be useful as an addendum, to allow neighboring agencies to contribute spatial lag effects and address autocorrelation concerns. Longitudinal (and therefore multi-level) models may be appropriate to account for year-specific rates. Large residuals / variance from that overall model are expected to be suggestive of different, but unknown or unspoken local policies. The residuals of the model themselves may be associated with differences in race-ethnic disparities in overall stop rates.

Assessing these stop type correlations with their respective hypothesized ecological correlate may be best done using three entirely separated models, or by using a multivariable, multivariate model that uses crash rates, crime incidence, and poverty rates to simultaneously model moving violation stops, subjective stops, and economic stops. Graphical approaches (e.g. regular map making) may be useful in considering whether there are unmodeled confounders or alternate theories of correlation that are not visible in the purely mathematical model.

Multivariable clustering techniques borrowed from data science may also be useful as an exploratory method to describe and name types of LEAs stop profiles, implying clusters of underlying stop policies. As example, one group of agencies may perform fewer safety stops, but more economic and subjective stops, have high racial disparities, and high crash incidents – suggesting this group may benefit from similar policy interventions to Fayetteville approach, reprioritizing safety stops and deprioritizing economic and subjective stops to attempt to address these patterns. Other agency profiles (to be determined) may suggest other interpretations and natural policy responses.

Specifically, I offer three high-level, but testable hypotheses to assess the link (or absence thereof) between of traffic stops and public health related outcomes, and a final model to explore the patterning of these hypotheses.

Hypothesis 1: safety stops reduce vehicle crash outcomes

First, I hypothesize (1) the rate of safety-related traffic stops should closely model vehicle crash rates, optionally weighted by crash severity, with possible effect measure modification by urban/rural status of the jurisdiction (from census data) or the distribution of road mileage per area in that jurisdiction. As mentioned prior as concerns the Deterrence Theory and highway safety research, a preliminary literature review to this hypothesis has not yet turned up conclusive evidence of LEA traffic stops affecting crash outcomes. Aim 2 seeks to add to that body of literature, but as is, variation between ostensibly safety-related stops and crash rates may represent underlying, latent LEA policies and practice patterns.

Hypothesis 2: economic stops mirror poverty rates

I further hypothesize that (2) the rate of economic (regulatory & equipment) stops should closely model poverty rates in that jurisdiction's residential or driving based denominators. That said, a framework that acknowledges the disproportionate impact of fines on low-income community would recommend against fining drivers for already being low-income (and thus, less able to address vehicle registration, insurance and maintenance issues) as adding financial insult to injury. In addition, the public health benefit to this practice seems suspect, while the financial harm to low-income communities seems demonstrable. In accordance with recommendations from recent years of organizing around this effort by community organizations, some agencies have elected to dramatically reduce their economic stops. With appropriate confounding control, we hypothesize that LEA variation in residuals to the association of economic stops and poverty rates suggest underlying, latent policies of policing poverty worth revealing. Following critiques from the US DOJ report on racial disparities in justice outcomes in Ferguson, if data on the rate of ticket-based local funding is available (e.g. from the administrative office of the courts), this may also be a useful predictor of agency specific economic stops.

Hypothesis 3: investigatory stops reduce crime rates

Lastly, I hypothesize that (3) the rate of subjective stops should closely model crime rates. This hypothesis is more challenging, given the nature of these stops. For instance, this stop category includes seatbelt stops. As previously mentioned, some LEAs followed public health's recommendation to pursue seatbelt stops as a means of reducing the injury severity of crashes. That said, seatbelts are an eminently usable pretextual reason for a stop, allowing perhaps more discretion on the part of the LEO stopping a driver. And as suggested results from Baumgartner (2019) seatbelt stops seem to be disproportionally employed with Black drivers. Some, but perhaps not all, of that difference may be due to differences in seat belt usage. But these stop disparities may represent neighborhood-specific application of this law, pretextual use of seatbelt stops for other unstated purposes, or perhaps widespread differences in seatbelt use by race-ethnicity. Variation between agencies may shed light on these hypotheses and more. Evidence of investigatory stops reducing crime rates seems lacking, and Aim 2 provides some evidence of the a related, opposite possibility: that large reductions in investigatory stops may have been associated with no increase in incident index crimes or violent crimes. It may be that only a very small and targeted subset of traffic stops have any crime effect, and the others could be reduced or stopped with little negative consequence from crime (and potentially positive consequences for community trust).

Taken together, these three traffic-stop-related percentages (along with, optionally, there optionally associated health outcomes) represent the explicit or implicit prioritization of certain health outcomes, and a kind of fingerprint for agencies

Model: Create an ecological model to explain variation against these hypotheses

Relevant ecological variables are expected to predict agency-specific stop profiles (of both total stop rates and stratified stop types). These include both injuries intended to be prevented by stops (e.g. violent crime and car crashes) and poverty demographics that regulatory and equipment stops are implicitly predicated upon. We describe a theoretical, idealized model of police traffic stop prioritization based on these public health priorities. However, other contextual data (to be determined) may be needed to describe variation against this ecological model, such as demographic representation on the force, local government structures, and proxies of local power (such as voting turn out). Variation may be in keeping with modern and prevalent critical theories of policing, such as the racial threat hypothesis that suggests policing on behalf of a majority may differ based on the relative sizes of racial and ethnic minorities or marginalized populations [71]. Theories of community power and policing may also suggest that LEAs may exert less control (e.g. lower stop rates) in communities that are empowered to influencing policing policies and resist police autonomy, regardless of related outcomes of traffic injuries or crime. Non-correlation can be as important as correlation in understanding these jurisdiction dynamics; as example, a community study of violent crime rates and police killing / death by legal intervention rates found them largely uncorrelated, suggesting alternate theories of violence by police are required for understanding variation between agencies [1].

8.4.3 List the multiple components of a traffic stop alongside their disparity tests

As described in Chapter 6, the Public Health Critical Race Praxis is a useful tool in critically examining traffic stops. That tool suggests conventional narratives around traffic stops are too limited, focusing on interpersonal bias exclusively, and ignoring the multi-level dynamics of law enforcement agency programs that drive and contextualize traffic stops. The PHCRP inspired figure from Chapter 6, while useful in contrasting narratives at multiple levels, does not model the interrelatedness of population disparities and agency program exposures and outcomes. Further, it does not provide tests for those disparities. For instance, different tests are appropriate when considering distribution of traffic stop locations at the agency level, individual officer bias, citations and warnings following stops, disparities in searches pursuant to a stop, contraband hit rates, and arrests after finding contraband. Related population and outcome may likewise reach beyond the measures used in this analysis (disparities in vehicle access, driving volume, and driving patterning) to consider income disparities, jail and prison incidence (and for what offenses? Drug-related?), crime incidents of multiple kinds, and motor vehicle crash data as covered in Aim 2.

While the previous section advocates for a specific series of tests involving crime, crash, and traffic stop rates by type, and while those hypotheses are under a particular framework, it is still a relatively small corner of the traffic stop picture. A more comprehensive, critical public health theory of traffic stops could convey the gamut of relevant and testable hypotheses. This would not only support a larger research plan but may have attached community actions to reduce disparities at each of these testable points. Taken together, these programs and dynamics might approach an actionable "theory of everything" for traffic stop programs from an anti-racist public health framework. Such a model would be best done captured visually and could be used for educating law enforcement about the levels and types of disparities attached to traffic stop questions.

8.4.4 Formally critique RTI STAR

As covered in the introduction and literature review, RTI STAR is a highly targeted test that ignores the multi-level, multi-agent dynamics of traffic stops in favor of a hyper-focused test of interpersonal bias based on the veil-of-darkness (VOD) test. It is used by agencies with little regard to its limitations. As cited previously, one chief used phrases such as "for racial profiling to occur, the p-value would have to be 0.05 or less", speaking broadly about their traffic stop program (Daily Tar Heel, 2018). Law enforcement administrators are not statisticians, and while many academic organizations have raised alarms about this sort of interpretation overreach, especially placing such a high value on p-value hypothesis testing, I believe it is ultimately the responsibility of academics to make every effort to limit inappropriate uses of our research. We simply do not live in a world where lightly documented tools for defending police agencies seeking to defend themselves against accusations of racial disproportionality won't be mis-used in this way.

Counterexamples may be more instructive than nuanced critiques of the math and a paper on theory, here. Such counter-example analysis could be designed as follows. First, (1) based on the limits of the VOD test, construct a set of theoretical situations where agencies could pass the VOD test but fail any reasonable lay consideration of disparities. These counterexamples might include: (A) individual officers "hiding" high disproportionality in agency average estimates; (B) specific patrol areas with high disproportionality; (C) changing underlying demographics and travel populations; (D) changing disparities by certain stop types, again hidden by aggregation; (E) "p-hacking" the test by including too-few traffic stops to get a low-enough p-value, etc. Many such examples exist – all they share is that they maintain the same agency-level disparities, however high, by whatever mechanism, before and after dark.

The second step would be to (2) intentional construct datasets to pass through the publicly available RTI STAR tool with $p>0.05$, but which a community group or police administrator would recognize as being patently problematic. The required dataset is relatively simple, so building these simulated datasets would not be particularly difficult. Theorizing a collection of counterexamples is likely the more difficult task.

An optional third step might be to (3) find specific agencies in the NC traffic stop dataset that best exemplify these theoretical examples, passing the test in some ways, but failing other reasonable tests. Some of this may require supplementing the datasets, e.g. since sub-agency traffic stop data is not available for most agencies, intentionally distributing traffic stops in certain pattern so as to fail a neighborhood test while passing the aggregate test.

8.4.5 Sub-agency analyses

Both agencies and community groups have an interest in considering traffic stop rates, disparities, and associations within jurisdictions, e.g. in neighborhoods or by road segments. The supplemental analysis for Aim 2 contains possible techniques for this sub-agency analysis and are explored in more detail in an Appendix. Some of this work may be worth extending into future manuscripts to demonstrate how to assess small-area disparities and design evidence-based traffic stop programs that are accountable to public health outcomes and community priorities. However, sub-agency analyses have unique challenges, not the least of which is the challenge of small area estimation of denominators. Still, sub-agency analyses enable otherwise impossible research aims, such as exploring the within-agency distributions of patrol patterns and the proximity relationships of traffic stops and public safety incidents.

8.5 **Areas for anti-racist action**

Beyond the continued research this dissertation implies, the results of these analyses suggest certain anti-racist action from community coalitions and engaged researchers. These include but are not limited to the following three focus areas:

8.5.1 End traffic stops that police poverty

Over 100 years ago, French author and philosophy Anatole France observed “the poor have to labor in the face of the majestic equality of the law, which forbids the rich as well as the poor to sleep under bridges, to beg in the streets, and to steal bread.[4]" In keeping with that observation, all drivers must maintain appropriate insurance, car registration, and vehicles in working condition regardless of socio-economic position, wealth, and income. However, few wealthy drivers would ever have these issues, so this "majestically equal" law is effectively policing poverty. More than that, when considered alongside associated fees, traffic stops for regulatory and equipment reasons create negative feedback loops further extracting wealth from already disadvantaged communities. Further dashboard / fact sheet work could highlight the raw number of these kinds of stops, or use linked administrative office of the courts data on fines to total the cost to low-income people of these programs.

8.5.2 Increased disparities and equity accountability infrastructure

Many of the possible traffic stop interventions, including the one explored in Aim 2, could be better served by collecting additional data to assist not only in program design and evaluation, but community accountability. Two additional data elements, at least, would be of particular use. First, unlike motor vehicle crashes and much crime data, the vast majority of agencies do not elect to collect data on the location of their traffic stop within their jurisdiction. This not only disallows agencies and researchers from considering the relationship of traffic stops to related public health outcomes, but also disallows analysis of neighborhood-specific patrol distribution patterns that informed communities might not consent to. Second, while some agencies have elected to capture the city and county of residence when pulling over drivers

[4] "La majestueuse égalité des lois, qui interdit au riche comme au pauvre de coucher sous les ponts, de mendier dans les rues et de voler du pain." Anatole France, The Red Lily, 1894.

(useful data when considering travel patterns), this is also not required on the form. At face value these seem like simple additions.

However, anecdotal discussion with lawyers and policy makers through community collaborations with the author suggest a hesitancy, in this political climate, to propose any changes to the SBI-122 form lest the state legislature, reminded of the existence of the program, do away with the existing reporting requirements. Many law enforcement administrators, in response to advocacy from community coalitions, have cited the cost of GPS programs. This is a difficult position to defend, as GPS tools are increasingly inexpensive and packaged on every mobile phone, and agencies are considering much more expensive hardware and storage packages for body and dash cameras. This data has to be input alongside the SBI-122 form, however, so vendor costs for integration (e.g. adding a field to a form) may be high. Perhaps using a free text field for nearest address or intersection, which would then retroactively be geocoded en masse, might enable agencies to do this.

However, data collection and analysis to drive change, including the way it's used in this dissertation, may be considered a "White method based on White logic" [149]. While sometimes useful for anti-racist ends, it is far from the only way to ensure accountability to community and achieve reductions in disparities. Other community tools, such as organizing more generally and police accountability bodies more specifically, are also important tools. Better representation in local government of those most impacted by policing strategies may be a formal response, but collective responses through non-government organizations and informal bodies also drive policy and accountability. Likely it will take a mix of many strategies to increase accountability.

8.5.3 Structural change: Fund policing alternatives / abolish policing as we know it.

While reducing or ending enforcement of certain kinds of traffic stops may be a viable stop gap strategy in some areas, it is far from a structural change to policing. Incremental redesign ignores the reality that many forms of and resources for policing have changed dramatically over time, and incremental change may not be a viable long-term change strategy even in places it is a worthwhile stop gap, harm reduction tactic. Additionally, in some jurisdictions this sort of incremental redesign may be more difficult than focusing on a deeper structural change from the beginning.

In light of this, some progressive researchers, activists, authors and community groups have outlined plans, some more realistic than others, to structurally redesign policing. These efforts are important to consider seriously for two reasons. First, incremental redesign may ignore other system-wide improvements that may operate better by diverting resources from policing strategies, however efficient within that sector, to non-police sectors that may be better able address public safety needs than enforcement strategies. For instance, while there has been interest in mental health crisis training for officers, no amount of training of officers may be more cost efficient or outcome effective in the long term than appropriately funding mental health services. Timely enforcement interventions for extremely unsafe behavior may be useful, but must consider other evidence-based strategies to prevent unsafe behavior through infrastructure and other means. Violence and the harms of substance abuse each have other strategies as well. Besides efficiency and efficacy, interventions must also consider equity: not all interventions that work at the population level carry as much collateral damage to marginalized populations. For instance, while the CDC acknowledges saturation patrols and costly fees for seatbelts reduce some negative health outcomes for motor vehicle crashes, these interventions do significant collateral damage on their own to the resources of low-income communities and may be more at risk of violating community trust, which has its own list of associated public health harms [34].

In contrast, recent scholarship on policing accountability and particularly scaling back of police activities have documented a potential “Ferguson effect,” where increased accountability and scaling back of services leads to an increase in negative public safety outcomes, such as violent crime. This scaling back of services may be intentional and explicit, or it may be due to cultural conflict within organizations, leading some individual officers to essentially strike while on the job, refusing to perform duties because of a lack of community trust or administrative support. However, while anecdotal discussion of the Aim 2 Fayetteville intervention did find cultural change was difficult and resulted, for a time, in many fewer traffic stops, it did not document any sort of Ferguson effect on measures of crime. In contrast, even with fewer stops, the efficiency of those stops to prevent traffic crashes may have increased with little to no measurable change in crime incident outcomes. This finding is in keeping with other studies that have acknowledged cultural challenges when scaling back enforcement efforts and increased accountability but found no Ferguson effect in their interventions [100,121]. Other studies that have investigated “de-policing” also acknowledge Ferguson effects are often confounded by population growth, racial segregation, lower levels of educational attainment, and poverty, and may be as much driven by community non-cooperation because of a lack of trust as a reduction in output by agencies [62]. Accordingly, a study of officers finds this Ferguson effect real in the attitudes of officers, but less so if those officers believe community legitimacy and trust are important factors for agencies [101].

It is essential that public health, and evidence-based, ethically responsibly policing, consider the most efficient, effective, and equitable interventions from all sectors, not simply grow enforcement activities in all directions while underfunding stronger interventions. This may be a difficult argument for self-serving law enforcement agencies that hope to grow without limit to agree with. However, agencies may find it is paradoxically in their best interest to scale back their activities to ones that are the most useful, most cost efficient, and most equitable. And regardless of the self-interest of any government agency, communities should have a fundamental right to representation and self-determination of policing strategies. Majority (and in some cases, powerful minority) rule that defines inequitable patrol and program priorities that target underrepresented populations within their jurisdictions will perpetually be challenged by legitimacy and trust concerns.

8.6 **Conclusion**

Baumgartner et al. [16]include a section titled “Why bother?” when considering statistical tests of traffic stop disparities, given the overwhelming evidence of race-ethnicity disparities in law enforcement related measures. This is a reasonable question, especially as communities in the United States, perhaps particularly communities of color, have known law enforcement programs, including traffic stop programs, operate with severe disparities for a long time. As reviewed previously, notable Black community research on traffic stops, centering stories and experiences along with data collection, lead to the publishing of the Green Book decades prior. Given this community knowledge, what is the role data-based studies such as these?

Baumgartner et al. list five reasons: (1) NC law established the program with the intention of assessing disparities, which for various reasons was not acted on; (2) knowing disparities exist is not the same as knowing their magnitude, especially when considering relative magnitude compared to other agencies, (3) specific measurement of disparities is required in order to assess trends, (4) measurement of disparities provides a framework to interrogate their components, and (5) measurement of disparities provides a framework to interrogate their causes.

However, quantification of disparities in the way done in this dissertation is not without its fundamental limitations and negative effects, beyond the analysis limitations covered previously. At our worst, public health studies that center data over community knowledge are may be predatory, further extractive value from communities in many forms, including the power of who gets to represent issues and resources such as funding for current and future studies. Providing quantitative evidence may further silence the voices of communities who only have their own direct experiences to speak from if those direct experiences are not viewed with equal or greater explanatory power than the results from mathematical models.

While in a world abstracted from inequity and the power differentials scientific knowledge could be generated for knowledge's sake without negative or differential consequences, we do not live in that world. As discussed in Chapter 6, quantitative research, perhaps particularly research on (vs. with, or by) marginalized populations, even if well intentioned, even if on issues of justice, may be particularly at risk of these negative consequences – consequences irrespective of intentionality. More practically, quantitative research like this may (but not necessarily will) be able to be used by communities for enhanced power for self-determination. Research may also be used to oppress, and even well-designed studies may have their limitations ignored or findings misinterpreted for harmful ends. This study does, and future studies should, acknowledge these dangers and help to hold up the direct experiences of communities for self-determination and control over their own environments, including enforcement actions in those environments.

APPENDIX 1: SBI-122 TRAFFIC STOP REPORT

TRAFFIC STOP REPORT

Agency Name | Date (Month/Day/Year) | Time

County of Stop | Officer ID Number

City of Stop

Part I

Initial Purpose of Traffic Stop *(check only one)*

- ☐ Checkpoint
- ☐ Driving While Impaired
- ☐ Investigation
- ☐ Other Motor Vehicle Violation
- ☐ Safe Movement Violation
- ☐ Seat Belt Violation
- ☐ Speed Limit Violation
- ☐ Stop Light / Sign Violation
- ☐ Vehicle Equipment Violation
- ☐ Vehicle Regulatory Violation

Vehicle Driver Information

Driver's Age ______ **Driver's Race** ☐ White ☐ Black ☐ Native American ☐ Asian ☐ Other

Driver's Sex ☐ Male ☐ Female

Driver's Ethnicity ☐ Non-Hispanic ☐ Hispanic *(Person of Mexican, Puerto Rican, Cuban, Central or South American, or other Spanish Culture)*

Enforcement Action Taken as a Result of the Traffic Stop *(check only one)*

- ☐ Citation Issued
- ☐ No Action Taken
- ☐ On-View Arrest → If arrest made, who was arrested?
  - ☐ Driver
  - ☐ Passenger(s)
- ☐ Verbal Warning
- ☐ Written Warning

Physical Resistance Encountered

Did Officer(s) encounter any physical resistance from Driver and/or Passenger(s)?	☐ Yes	☐ No
Did Officer(s) engage in the use of force against the Driver and/or Passenger(s)?	☐ Yes	☐ No
Did injuries occur to the Officer(s) as a result of the stop?	☐ Yes	☐ No
Did injuries occur to the Driver as a result of the stop?	☐ Yes	☐ No
Did injuries occur to the Passenger(s) as a result of the stop?	☐ Yes	☐ No

Vehicle/Driver/Passenger(s) Search

Was a search initiated subsequent to the traffic stop? ☐ Yes* ☐ No

***If search was initiated, complete Part II**

SBI-122 (Rev. 12/09)

Traffic Stop Report

Part II

Type of Search *(check only one)*

☐ Consent ☐ Search Warrant ☐ Probable Cause ☐ Search Incident to Arrest ☐ Protective Frisk

Basis for Search

☐ Erratic/Suspicious Behavior ☐ Observation of Suspected Contraband ☐ Suspicious Movement

☐ Informant's Tip ☐ Other Official Information ☐ Witness Observation

Person(s)/Vehicle Searched

Was the Vehicle Searched? ☐ Yes ☐ No
Was the Driver Searched? ☐ Yes ☐ No
Was a Passenger(s) Searched? ☐ Yes ☐ No
Were the Personal Effects of the Driver and/or Passenger(s) Searched? ☐ Yes ☐ No

Identify the sex, race, and ethnicity of each passenger searched

	Age	Sex		Race					Ethnicity	
		Male	Female	White	Black	Native American	Asian	Other	Hispanic	Non-Hispanic
Passenger 1										
Passenger 2										
Passenger 3										
Passenger 4										

Contraband Found

Contraband found as a result of the search: ☐ None <u>**OR** complete the following</u>:

☐ Drugs ____ Ounces ____ Pound ____ Dosages ____ Grams ____ Kilos

☐ Alcohol ____ Pints ____ Gallon

☐ Money ____ Dollar Amount

☐ Weapons ____ Number of Weapons

☐ Other ____ Dollar Amount

Property Seized

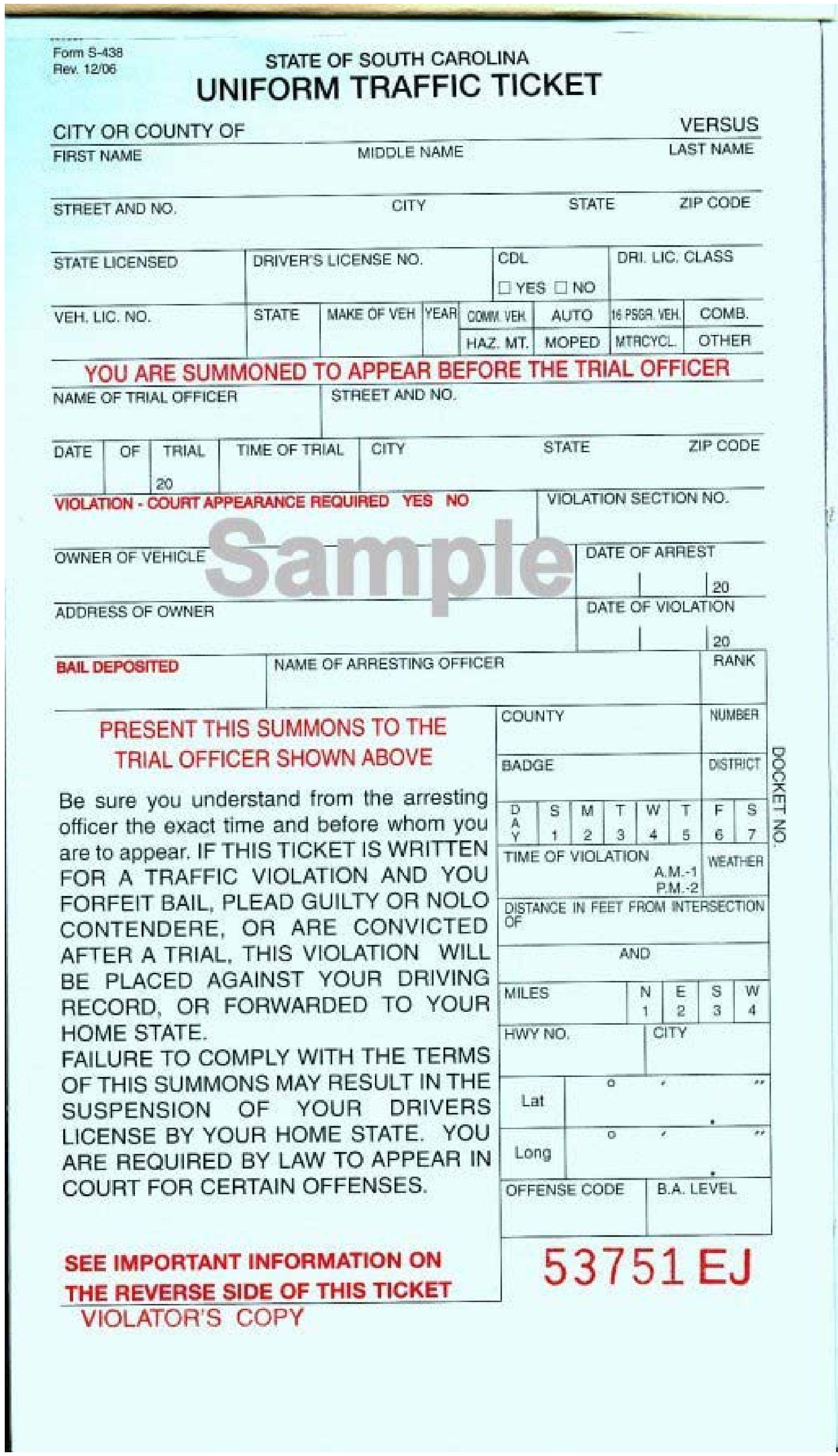
Property seized as a result of the search: ☐ None <u>**OR** complete the following</u>:

☐ Motor Vehicle ☐ Personal Property ☐ Other Property

Office Use Only	Date	Initials
Reviewed		
Entered		

SBI-122 (Rev. 12/09)

APPENDIX 2: SOUTH CAROLINA TRAFFIC TICKET FORM S-438



Form S-438
Rev. 12/06

STATE OF SOUTH CAROLINA
UNIFORM TRAFFIC TICKET

CITY OR COUNTY OF ... VERSUS

FIRST NAME | MIDDLE NAME | LAST NAME

STREET AND NO. | CITY | STATE | ZIP CODE

STATE LICENSED | DRIVER'S LICENSE NO. | CDL ☐ YES ☐ NO | DRI. LIC. CLASS

VEH. LIC. NO. | STATE | MAKE OF VEH | YEAR | COMM. VEH. | AUTO | 16 PSGR. VEH. | COMB. | HAZ. MT. | MOPED | MTRCYCL. | OTHER

YOU ARE SUMMONED TO APPEAR BEFORE THE TRIAL OFFICER

NAME OF TRIAL OFFICER | STREET AND NO.

DATE OF TRIAL 20 | TIME OF TRIAL | CITY | STATE | ZIP CODE

VIOLATION - COURT APPEARANCE REQUIRED YES NO | VIOLATION SECTION NO.

Sample

OWNER OF VEHICLE | DATE OF ARREST 20

ADDRESS OF OWNER | DATE OF VIOLATION 20

BAIL DEPOSITED | NAME OF ARRESTING OFFICER | RANK

PRESENT THIS SUMMONS TO THE TRIAL OFFICER SHOWN ABOVE

Be sure you understand from the arresting officer the exact time and before whom you are to appear. IF THIS TICKET IS WRITTEN FOR A TRAFFIC VIOLATION AND YOU FORFEIT BAIL, PLEAD GUILTY OR NOLO CONTENDERE, OR ARE CONVICTED AFTER A TRIAL, THIS VIOLATION WILL BE PLACED AGAINST YOUR DRIVING RECORD, OR FORWARDED TO YOUR HOME STATE.
FAILURE TO COMPLY WITH THE TERMS OF THIS SUMMONS MAY RESULT IN THE SUSPENSION OF YOUR DRIVERS LICENSE BY YOUR HOME STATE. YOU ARE REQUIRED BY LAW TO APPEAR IN COURT FOR CERTAIN OFFENSES.

COUNTY | NUMBER

BADGE | DISTRICT

DOCKET NO.

DAY: S 1 | M 2 | T 3 | W 4 | T 5 | F 6 | S 7

TIME OF VIOLATION A.M.-1 P.M.-2 | WEATHER

DISTANCE IN FEET FROM INTERSECTION OF ... AND

MILES | N 1 | E 2 | S 3 | W 4

HWY NO. | CITY

Lat ° ' "

Long ° ' "

OFFENSE CODE | B.A. LEVEL

SEE IMPORTANT INFORMATION ON THE REVERSE SIDE OF THIS TICKET

VIOLATOR'S COPY

53751 EJ

NOTICE

THE PRIMARY AIM OF TRAFFIC LAW ENFORCEMENT IS TO REDUCE TRAFFIC ACCIDENTS, INJURIES AND DEATHS THROUGH FAIR, IMPARTIAL, AND REASONABLE ENFORCEMENT OF TRAFFIC LAWS.

You must settle the case hereby made against you in one of three ways.

1) You must appear in court at the appointed time with your South Carolina drivers license if indicated in the violation block on the front side of this summons.
2) You may post a cash bond with the appropriate trial officer's office prior to the assigned date of trial. If you decide to mail in your bond rather than go to the office, MAIL MONEY ORDER, CASHIER'S CHECK OR CERTIFIED CHECK DIRECTLY TO THE TRIAL OFFICER'S OFFICE BEFORE WHOM YOU ARE SUMMONED TO APPEAR. The trial officer's name and address is shown on the front side of this summons. DO NOT MAIL CASH OR PERSONAL CHECK. Be sure to enclose with the bond the arresting officer's name and the summons number.
3) You may appear in court on the assigned date and time and have a trial conducted by the trial officer.

THE POSTING OF BOND BEFORE YOUR ASSIGNED TRIAL DATE IN NO WAY AFFECTS YOUR RIGHT TO HAVE A FAIR TRIAL BY THE JUDGE OR, IF YOU MAKE A WRITTEN REQUEST BEFORE YOUR SCHEDULED TRIAL, BY JURY.

However, if you are NOT required to appear in court on the assigned trial date and have previously posted bond and do not appear on the trial date, your bond may be forfeited unless the judge has agreed to have your case heard at another time.

If you fail to post bond or personally appear in court on the assigned trial date, your home state's Motor Vehicle Division will be notified to suspend your license until you have cleared this matter with the trial court. Additionally, a willful failure to appear or post bond is punishable as a separate offense by a fine of up to $200.00 or imprisonment of up to 30 days.

APPENDIX 3: FAYETTEVILLE SUB-AGENCY ANALYSIS

Container-based, neighborhood-level analysis

Administrative boundaries sometimes mirror meaningfully different activity spaces, sometimes with both separate formal policies, practice and demographic patterns, as well as separate informal cultures, landscapes and intra-area dynamics. When administrative boundaries map well to locally recognized neighborhoods, this neighborhood-specific analysis can be a useful method for driving local policy conversation.

Neighborhood level analysis can help ground-truth the implementation of the Fayetteville intervention and identify areas for further focus. For instance, below, the percent black in the block group population (as dots for >60% and >80% black) is layered over the percent of stops that were income-related, demonstrating a remaining neighborhood-level association between demographics and the proportion of stops by type. Should Fayetteville, like NC as a state, have racial disparities by income, then income demographics may partly describe this phenomenon. Regardless, Black residents experience of policing in their neighborhoods is associated with proportionally more regulatory and equipment related stops than whiter neighborhoods.

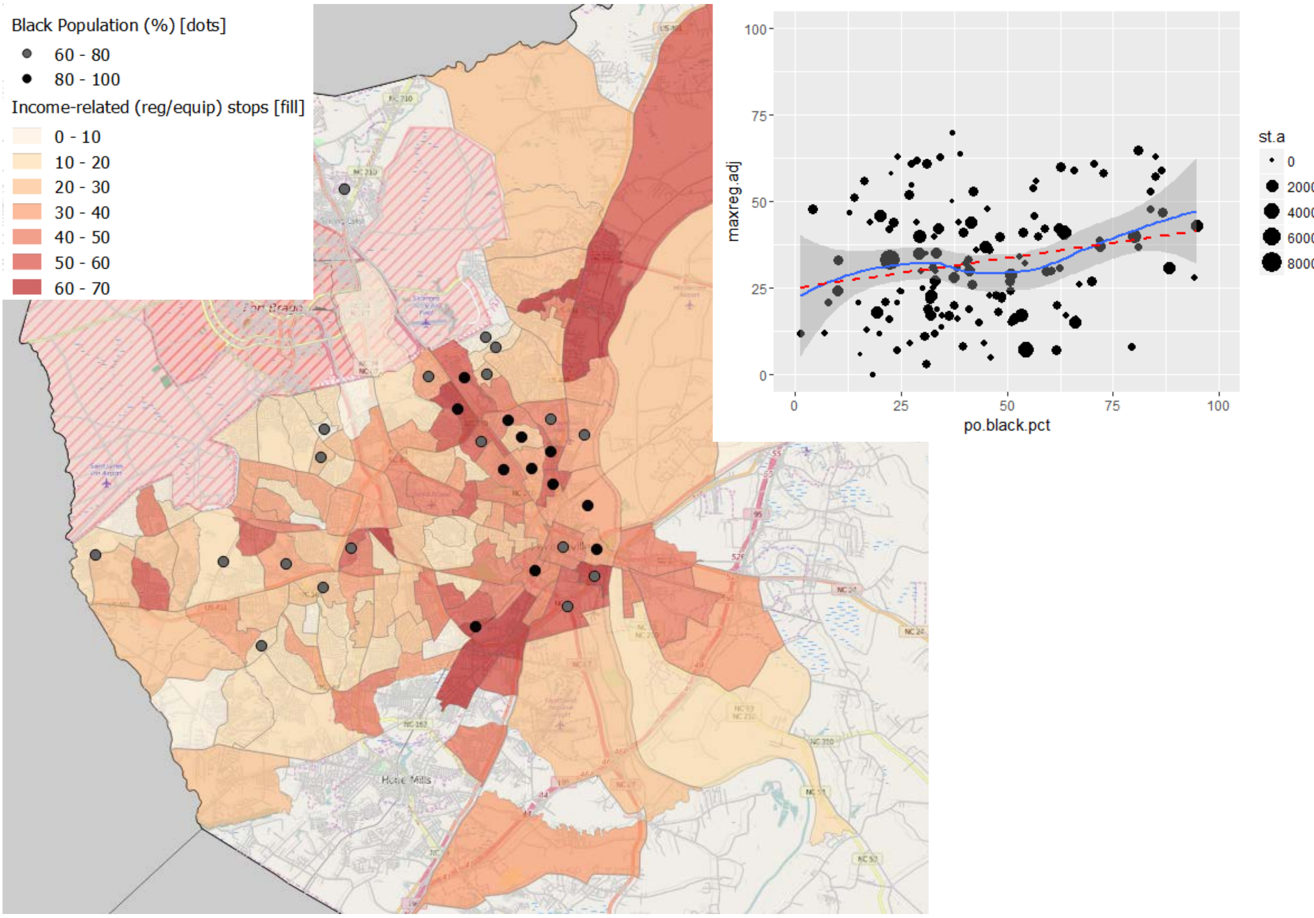


Figure A3-1 Neighborhood-specific percent income related stops and % black population.

Preliminary analysis suggests higher percent Black communities seem to be largely the same ones where a high percent of regulatory stops occur. Each 5% increase in neighborhood percent black corresponds to an additional 1% increase in the percent of people pulled over for regulatory reasons (above right).

Specific, neighborhood-local stories can be useful in ground truthing the model and intervention action. Below are four block groups and their point-level crashes and stops from 2013-2015. The mall (bottom) saw the number two stop increased (over 1,000 more stops a year), and was the number 1 injury area for Fayetteville by count, with three times the injuries of any other block group. The top two blocks sandwich the American Expressway and Rt. 24 at

Santa Fe/Shaw Road, which each saw a high stop increase, and had high injuries relative to other block groups. On the right an 80% Black community saw the largest decrease in stops, with relatively few traffic crashes to drive stops. These provide important anecdotal evidence of intentional clustering of traffic stops to high crash areas.

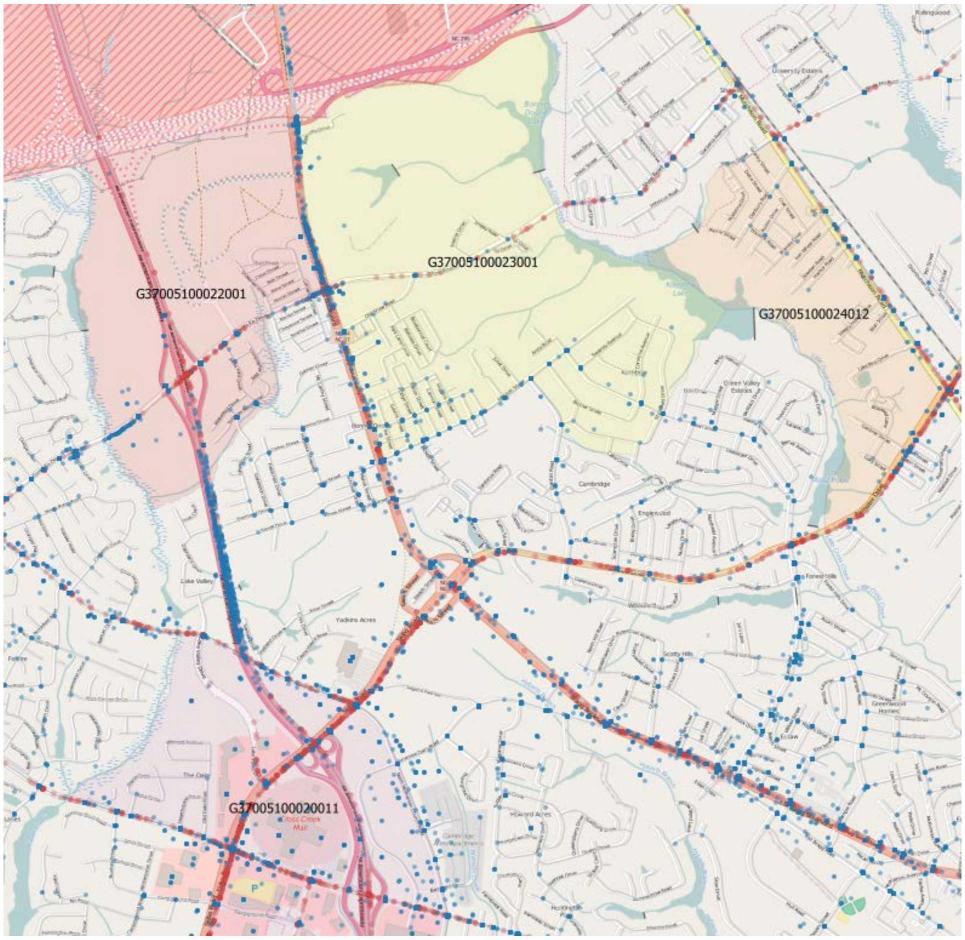


Figure A3-2 Four block-group changes in Fayetteville Police Department stop prioritization, 2013 to 2015.

Bottom: Mall: #2 stop increase (+1000/y), #1 injuries. 3x the injuries of everywhere else. Top two: All American Expressway and Rt. 24 at Santa Fe/Shaw Rd. #1 stop increase, #7 injury and #3 stop increase, #24 injury. Right: #1 stop decrease, #66 injury. 80% B/AA. Note areal vs. street placement.

Area-level GLM, GWR, and autocorrelation adjustments can be useful techniques in describing the spatial association between variables of interest. Below are the residuals to a linear model, spatial lag model, and GWR model of change in stops against change in crashes over the time-period of interest, along with the beta coefficients of GWR model.

Overall, regions with more accidents did see an increase in police stops compared to those with fewer accidents. Broadly, for every additional accident each year, the PD stopped another 2 drivers in 2015 vs. 2013. Though this was truer for some areas than others, and there were some exceptions: in some areas a one more traffic accident was associated with 5.5 more police stops. In other areas a one more traffic accident was associated with one fewer police stop. Overall, for every additional accident in 2013-2015, Fay PD stopped an additional 0.6 (0.1 standard error) drivers comparing 2015 and 2013. This sub-jurisdiction analysis then, in addition to be a useful check on neighborhood anecdotal stories, helps to validate the implementation.

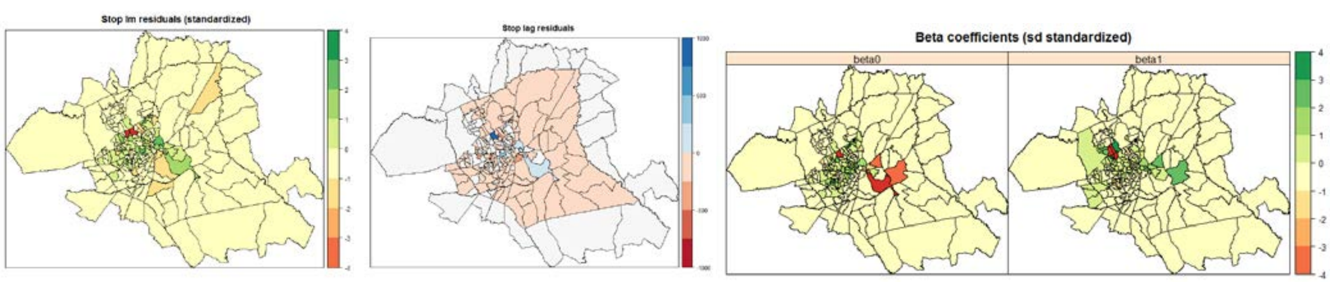


Figure A3-3. Fayetteville block-group residuals and beta coefficients used a crude linear model, spatial lag model, and explored using geographically weighted regression.

As useful as container-based analysis sometimes is, administrative boundaries have known limitations. Boundaries typically end on roads centerlines, meaning traffic stops on the same road might be apportioned to one or the other facing "neighborhood." Encapsulated areas do not always have local meaning beyond an administrative one, arbitrarily dividing or grouping more useful sub-divisions and biasing spatial areal effects toward the null. Therefore, we supplement this analysis with small-area, surface-based analysis that avoids the use of administrative boundaries.

Small-area surface-based analysis

Fundamental to the any area-based geospatial analysis are decisions about aggregation units. In the above preliminary analysis, I used block groups to balance block-level precision with capturing neighborhood boundaries. But the modifiable areal unit problem (Openshaw, 1984) acknowledges that different aggregation of sub-units of a spatial analysis can produce different results, making analysis prone to sometimes arbitrary delineations of space; gerrymandering is a modern instance of this.

Therefore, quarter-mile square aggregations were chosen based on two methods. First, qualitatively, I examined the natural clustering of traffic stops and crashes visually against map with a focus on intersections and road sections. This suggested a quarter mile block balanced capturing of activity at intersections with a natural, smooth fall-off effect theorized from the underlying continuous driving process. Second, I modeled the count of events and variance of the underlying space-time field that produced our given stop and crash distribution by a more complex and modern quantitative method, Bayesian Maximum Entropy (see Power for details), that validated that ¼ mile aggregation choice.

The below heat map set describes traffic stops, crashes, incidents, and offers a mechanism for choosing small-area prioritization for traffic safety stop or discretionary stop law enforcement.

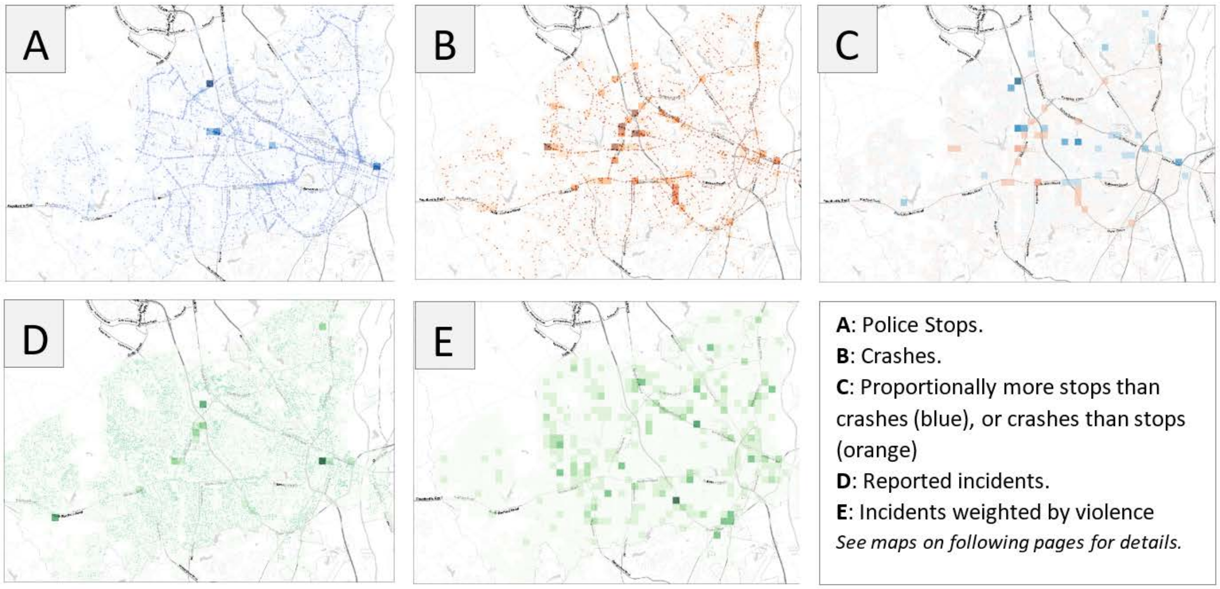


Figure A3-4. Police stops for moving violations (A), traffic crashes (B), raster subtraction of A and B (C), reported incidents (D), and injury severity weighted incidents (E).

Police vehicle safety-related stops in 2016 are shown in blue in map A, with the grid showing area of higher percentage of all stops. This includes driving impaired, stop signs, speed limit, and safe movement stops. Note that, as a reminder to the theoretical groupings offered above, this does not include seatbelt-stops – though highly safety related in theory, since it often has high racial disparity in practice, it may be highly subjectively used. A case could be made to include seat belt stops as a means through which fatalities are reduced; however, the literature on that effect seems sparse.

Vehicle crashes (unweighted) is in map B, with 2016 crashes mapped as points, with a grid showing high % distribution of those points in a quarter square mile area. These are unweighted for severity of injury.

Difference in police vehicle safety stop percent and vehicle crash percent is in map C. This map can be used to see areas where police stop proportionally more drivers for safety reasons than the distribution of vehicle crashes would indicate. Though preliminary, maps like this could be used for specific stop types to fine-tune deployment patterns. This is a simple map, subtracting the percent of stops from crashes, but more advanced analysis is possible, so that "neighbor" grid points can contribute to each other's calculations using a spatial lag model on this gridded data. This map should be considered a proof of concept, since grid-based spatial subtraction may ignore real scenarios where crashes happen in intersections, but drivers pulled over may happen just outside of that intersections' grid. Note that the central question to this dissertation, the estimation of smaller, demographic-specific driving denominators is still a challenge here, but I hypothesize that safety stops and vehicle crashes may ultimately rely on a very similar driving denominator, suggesting they may be collinear enough that differences in the percent of stops and percent of vehicle crashes, and residuals to that relationship, may be meaningful enough for action.

Police incidents (weighted by coded injury severity / risk of injury) are in maps D and E. Over one hundred sixty incident categories are either combined by sum in map D, or in map E (tentatively) scored between a 1 and 5, with 1 as low/no crime, no risk of injury, 2 as property & low-risk drug crime, 3 as personal safety endangered, could escalate to assault, 4 as actual assault/injury, 5 as homicide/suicide fatality. Scores 1:5 are then ranked on a log scale roughly mirroring quality-adjusted years of life lost: lowest risk as 0.01, low-risk as 0.1, danger as 1, assault as 5, death as 50. Though categorizations of severe outcomes are obvious, other incident types are not, and community and police score may them differently. However, it is sensible that a homicide should not be worth the same in a model as a "suspicious person" (which, note, is not illegal). A homicide may be worth 50, an assault 5, suspicious person 0.1, and a cat in a tree is 0.01 – roughly mirroring quality adjusted years of life lost injury framework. These sorts of weights could be determined through a community input process, and are not without controversy (e.g., is coding sex work / prostitution high more likely to create a more unsafe or safe environment for sex workers by increased scrutiny?) but as is I have leaned on my own QALY-informed sensibilities.

Ultimately, **incidents could be combined with crashes, both weighted by injury severity**, to produce a small-area injury index or rate to help drive micro-patrol decisions.

In choosing a method for small-area modeling, **interpretability is paramount**. Police chiefs directing officers and community members understanding their community require evidence to be translated and actionable. I've tested some of these maps with police chiefs and community groups, and heatmaps like these seem to have some familiarity. A reasonable community member or police chief might ask: can we use injury maps like this to direct policing for maximum public health impact and minimize racial disparities? However, GLM/GWR may be useful for estimating relationships over time, allowing spatial lag effects, and identifying residuals worth exploring.

Displaying the individual data points for all traffic stops alone creates limits in interpretation, and modeling as a spatial point process against other point processes is challenging. One of the central challenges to viable spatial model results given large amounts of point-level data is the selection of aggregation size for small areas. This selection is non-trivial, and the incorrect selection can have negative consequences for model results, limiting the power to model and estimate relationships over time and space.

Aggregating this point data in some way, such as on a grid by count or percent of all stops, both benefits interpretation and simplifies models. Interpolation is the process of smoothing these aggregated point estimates into a surface. Interpolation methods such as inverse distance weighting (IDW) rely deterministically on a set number of points, and though widely used and simple to implement, are subject to limitations caused by incorrect selection of too-large or too-small aggregation regions. If aggregation grid is too large, covariance of the subsequent grid centroids will be very small, representing the lack of resolution in this continuous phenomenon. If the grid centroids are too small, data will be sparse, and covariance between the either infrequent or spatially-sparse points will be small. Further, IDW techniques do not model variance, making confidence intervals unavailable.

Bayesian Maximum Entropy (BME) framework

The **Bayesian Maximum Entropy (BME) framework** has been used to better model space-time processes in public health (Hampton, et al., 2011), including traffic-related public health outcomes like pedestrian injuries (Fox et al., 2015), and can directly inform intervention and prevention (Law, 2004). BME allows for the modeling of a global space-time mean trend (typically either zero, constant or linear), where in the non-zero cases the model then covariance models then describe the residual over than mean trend. During the process of integrating the global covariance models and mean trend with local data for estimation and variance, BME uses both maximum distance from the point of estimation (conceptualized as a limiting space-time cylindrical window defined using a coefficient to weight space and time to producing single space-time distance) and a minimum number of informative space-time points for local kriging. Like the global space-time trend selection, that local kriging window can be one of three kinds:

simple kriging (regressing to a mean of zero), ordinary kriging (modeling against a non-zero constant mean) or universal (modeling against a local linear space-time trend).

BME provides multiple improvements over a deterministic IDW approach that may benefit questions of traffic stops and public health outcomes. We focus on two of which here, since they benefit efforts to validate the space-time point aggregation window in Aim 4: (1) integration and interpretation of space-time covariance modeling and (2) using the modeled estimate's variance to validate selection of the space-time aggregation grid. These grids inform heatmaps, effectively raster grid choropleths, and are often used to visualize traffic- and LEA-related outcomes. Both the aggregation grid and covariance structure together, however, can inform space-time exposure-outcome association model choices such as space-time lags of the effect of traffic stops on crash prevention. BME relies on first modeling space-time covariance, and allows both deterministic, known, single data points (hard data) and probabilistic distributions of potentially known or inexact data (soft data) in space and time. Covariance in space-time describes the diminishing similarity of each point through time and each time point through space.

Though only used to validating binning choice, this is a novel and generalizable utilization of the BME framework. Building on a strong theoretical foundation, this additional use case can help spatial analysts making binning choices in other public health settings whether using a census of cases (such as in this case) or the more typical sample of values (e.g. environmental air or water sampling or sampling costly human lab results).

176,740 traffic stops by the Fayetteville Police Department (FPD) from 2013 to 2016 were gathered from FPD's data administrators and geocoded at the point-level. The individual geocoded stops were aggregated into a percent of all stops during that time-period. Both because of large changes in the number of stops per year (31,361 in 2013; 61,734 in 2016) and because of software limitations, grid points with zero stops were removed from the analysis; to account for this, the global mean trend was left at zero, though local mean trends were modeled as a constant (ordinary kriging). Data was represented as percent values ranging theoretically from 0 to 100, but in practice, after having dropped zeros and with a large distribution, empirically ranged from greater than zero to under three percent for any grid cell, with time represented as the number of months since December 2012 (January 2013 as month 1 to October 2016 as month 46). Percent data was log-transformed to better fit a normal gaussian distribution.

I chose an aggregation space- and time-span of ¼ mile and 1 month based on a visual exploration of the space- and time-scale that captures, without fracturing, effects at intersections and along roads. We therefore chose the grid to prioritize communicability and balance small-number issues but will test BME's ability to assess this time and space aggregation choice statistically.

We modeled the space-time covariance of this ¼ mile, monthly percent data as a homogenous / stationary simple random field (SRF) in space-time. For such a SRF, the mean trend is constant and its covariance structure for c is modeled as a function of only r and τ, the spatial and temporal distance, respectively, between each point and its neighbors. For a two-part covariance structure,

$$c(r,\ \tau) = c_{01}\mathrm{e}^{-\frac{3\mathrm{r}}{\mathrm{a}_{\mathrm{r}1}}}\mathrm{e}^{-\frac{3\tau}{\mathrm{a}_{\tau 1}}} + \mathrm{c}_{02}\mathrm{e}^{-\frac{3\mathrm{r}}{\mathrm{a}_{r2}}}\mathrm{e}^{-\frac{3\tau}{a_{\tau 2}}}$$

Equation A3-1: General two-part covariance structure of a space-time SRF.

where (1) c_{01} and c_{02} add up to the total c_0, the modeled constant variance of each point (e.g., where $r = 0$ and $\tau = 0$, called the "sill") and (2) $a_{r*}$and $a_{\tau *}$ are the spatial and temporal ranges, respectively, at which 95% of their associated covariance structure is lost. This two-part structure provides a balance of flexibility and simplicity, in that both space and time can have both a short and a long covariance component, representing a process that both changes in space-time in the short-term (e.g. in a small number of units of space-time) as well as longer-term covariance describing a standing or slowly diminishing baseline over space and time.

The covariance structure for FPD traffic stops, all together and stratified by stop group, and traffic crashes are in the table below. To introduce this table, we first describe in detail the results for all FPD stops together (though the table also describes their type-specific covariance structure). Its smoothed, observed covariance structure is below (Figure 18), with a blue line representing the modeled relationship (Equation 2). These specific results are interpreted in following paragraphs after the concepts are described.

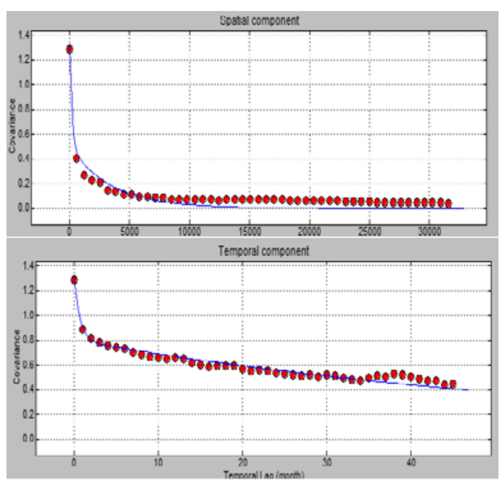


Figure A3-5. Observed (red dots) and modeled (blue line) covariance relationships for all (not disaggregated into groups) FPD traffic stops, 2013-2016.

$$c(r,\ \tau) = c_{01}e^{-\frac{3r}{a_{r1}}}e^{-\frac{3\tau}{a_{\tau 1}}} + c_{02}e^{-\frac{3r}{a_{r2}}}e^{-\frac{3\tau}{a_{\tau 2}}},\ c_0 = 1.3$$

$$c_{01} = 0.8,\ a_{r1} = 528\ ft\ (0.1\ mi),\ a_{\tau 1} = 200\ mo$$

$$c_{02} = 0.5,\ a_{r2} = 10560\ ft\ (2\ mi),\ a_{\tau 2} = 5\ mo$$

Equation A3-2: Covariance model for all (not disaggregated into groups) FPD traffic stops, 2013-2016.

	Covariance Structure						
Data Source	c_0	c_{01} (% c_0)	a_{r1}	a_{t1}	c_{02} (% c_0)	a_{r2}	a_{t2}
All Stops	1.29	0.75 (62%)	528 ft (0.1 mi)	250 mo	0.54 (38%)	10560 ft (2 mi)	5 mo
Safety	1.29	0.90 (70%)	2640 ft (0.5 mi)	100 mo	0.39 (30%)	79200 ft (15 mi)	3 mo
Economic	0.71	0.44 (62%)	528 ft (0.1 mi)	250 mo	0.27 (38%)	10560 ft (2 mi)	1.5 mo
Pretextual	0.29	0.19 (66%)	2640 ft (0.5 mi)	2 mo	0.10 (34%)	211200 ft (40 mi)	50 mo
Crashes	0.52	0.35 (71%)	528 ft (0.1 mi)	250 mo	0.17 (29%)	10560 ft (2 mi)	1.5 mo

Table A3-1. Covariance Structure of Fayetteville Traffic Stops, 2013-2016 and Traffic Crashes, 2016

Again, often the focus of BME modeling is in providing improved estimates of a space-time process by taking advantage of space-time covariance. In this case, we focus on the variance instead of the estimate to describe whether our aggregation strategy (1/4 mile grid, aggregated monthly) retains some covariance properties of its underlying continuous process, suggesting also we could interpolate between these aggregation centroids informatively. In short, we look to see (Figure 19) whether the modeled variance between known space-time points (blue with zero variance, representing the centroids of our estimation grid) increases to the maximum variance sill (red, in the below graph, at 1.29) between those points. If so, our covariance structure is insufficient to support interpolation, suggesting likely a too-large or too-small aggregation and a loss of covariance. The red max-variance field represents no known data at those point and estimation points not near enough to known data in space-time to benefit from our modeled covariance structure. Lighter dots of lower variance in red field represent points nearby in time influencing our estimate for this month (November, 2016). Examining a small slice, less than a mile by four miles, demonstrates these three variance scenarios: stop at this space time (blue dot with yellow kriging island), stops at that point at a previous time (yellow dot on red field), and no recent stops at all (red field) of the BME estimation variance (Figure 2) demonstrates that ¼ mile, 1-month aggregation is successful in capturing enough covariance to have informative interpolation between aggregation points, should IDW be used to smooth these aggregation points into a denser heatmap.

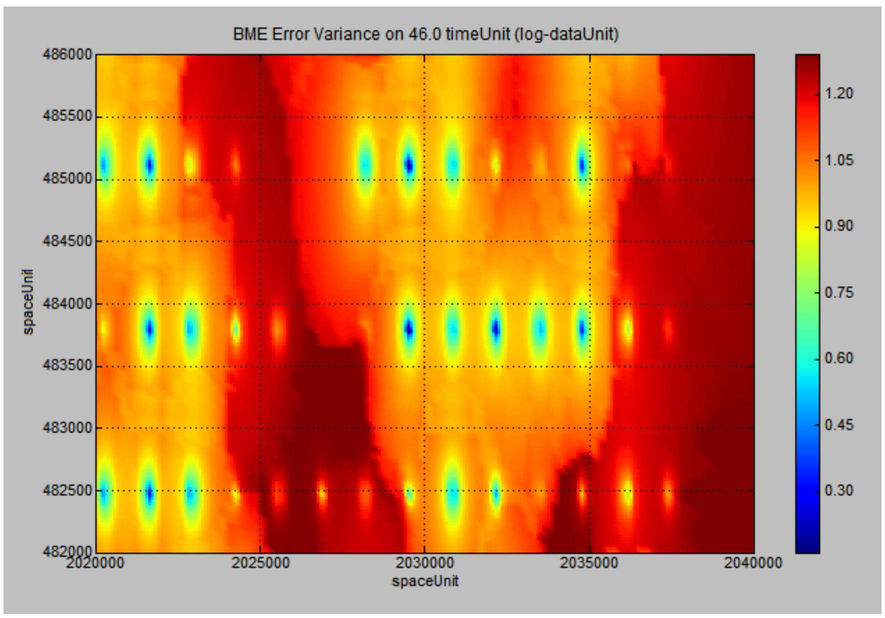


Figure A3-6. Subset of BME error variance map, Fayetteville

Returning to the table of model results for this ¼ mile, month grid method, we see that different types of traffic stops have different covariance structures. Comparing safety stops to economic stops, roughly 2/3 of the covariance is described by a short-distance (0.5 and 0.1 miles), long-time (100 and 250 months) structure and 1/3 of the covariance described by a longer-distance (15 and 2 miles) and short-time (3 and 1.5 months) structure. This corresponds to safety stops being more similar in both the short and long-term than economic stops, perhaps representing that traffic stops targeting larger areas. However, economic stops were more stable over a longer period of time, representing little long-term change in the distribution of those stops compared to the safety stops, an expected finding as the 2013-2016 intervention by FPD was to concentrate stops in higher crash areas, effectively changing the distribution across Fayetteville over the study period. Economic stops had the shortest long-term temporal range of 1.5 months, representing that the distribution of economic stops could change almost entirely month to month, suggesting their subjectivity and use for short-term neighborhood-level intents or department ticket / funding initiatives.

The pretextual stop covariance structure was similar to safety stops in the spatial component, in that 2/3 of the covariance was described by a short-term (0.5 miles) component and 1/3 by a longer-term component (though pretextual stops spatial lag was 40 miles, suggesting its flatter spatial surface than safety stops. The time covariance structure was different, however, with the 2/3 of the pretextual covariance distributed in the short term (2 months) instead of long-term like safety and economic. This reversal may represent their subjective nature, as 95% of the larger (66%) covariance component is lost over just 2 months instead of 100-250 months in the case of safety and economic stops. Their overall sill variance was also low, at 0.29 compared to 1.29 and 0.71 for safety and economic stops, again quantifying their subjectivity.

For comparison, traffic crash covariance was distributed in a structure that shared most similarity with economic stops: similarly, around 2/3 of covariance was in short-distance and long-time structures, with a long-time and space structures making up the remainder. One hypothesis for this might be that safety stops were known to be more human manipulated over this same period, and perhaps economic reasons stops are distributed more similarly over the driving surface in a similar the way that crashes are. In effect, this makes crashes, perhaps like economic stops, a more space-time random process, with less of a constant baseline trend.

These covariance structures have implications for understanding spatial lags in models that describe ideal causal effects between traffic stop types and consequent public health outcomes. Safety stops should ideally reduce traffic pedestrian crashes. Theoretically, according to police rationale, pretextual stops may have some reducing effect on crime, though this is unclear and community collateral damage to trust may be high. It is unclear what proximate public health consequence economic stops prevent, though it may be meaningful to explore whether their distribution mirrors the distribution of economic distress as measured by household adjusted gross income or percent below the federal poverty standard.

BME kriging variance gives some visualize interpretation to the extent of covariance fit across space and therefore the appropriateness of interpolating a heat map for interpretation purposes. In this case, in Fayetteville and with these types of traffic stop, ¼ mile, 1-month aggregation grid seems to be both small and large enough to capture some covariance for modeling purposes. A sensitivity analysis, expanding and shrinking the grid, or translating the grids across space may help bolster these results by testing how sensitive covariance is to variation in grid alignment and size. In the future, exploring these relationships using stop counts with a Poisson-based kriging instead of stop percent may be a more appropriate exposure construct for public health outcomes like prevention of crashes and LEA incidents.

In conclusion, BME kriging variance suggests that ¼ mile aggregation may be a balanced choice for modeling the underlying point process of both crashes and stops, sufficiently “powered” to model small-area effect estimates at any given location without having estimates between known points fall off to the maximum variance observed with no data at all.

REFERENCES

1. 2017 Police Violence Report. https://policeviolencereport.org/. Accessed July 14, 2019.

2. A Brief History of the Traffic Stop | Observer. https://observer.com/2016/07/a-brief-history-of-the-traffic-stop-or-how-the-car-created-the-police-state/. Accessed June 11, 2019.

3. Abadie A, Diamond A, Hainmueller J. Synth : An R Package for Synthetic Control Methods in Comparative Case Studies. *J Stat Softw*. 2011;42(13). doi:10.18637/jss.v042.i13.

4. Abadie A, Diamond A, Hainmueller J. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *J Am Stat Assoc*. 2010;105(490):493-505. doi:10.1198/jasa.2009.ap08746.

5. Abadie A, Gardeazabal J. The Economic Costs of Conflict: A Case Study of the Basque Country. *Am Econ Rev*. 2003;93(1):113-132. doi:10.1257/000282803321455188.

6. About Us - The Movement for Black Lives. https://policy.m4bl.org/about/. Accessed September 5, 2016.

7. Adanu EK, Smith R, Powell L, Jones S. Multilevel analysis of the role of human factors in regional disparities in crash outcomes. *Accid Anal Prev*. 2017;109:10-17. doi:10.1016/j.aap.2017.09.022.

8. Alan Blinder. Michael Slager, Officer in Walter Scott Shooting, Gets 20-Year Sentence - The New York Times. *New York Times*. https://www.nytimes.com/2017/12/07/us/michael-slager-sentence-walter-scott.html. Published December 7, 2017. Accessed May 27, 2019.

9. Alexander M. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. The New Press; 2012.

10. American Civil Liberties Union of Illinois. *CPD Traffic Stops and Resulting Searches in 2013*.; 2014. https://www.aclu-il.org/sites/default/files/wp-content/uploads/2014/12/Report-re-CPD-traffic-stops-in-2013.pdf. Accessed July 3, 2019.

11. Amrhein V, Greenland S, McShane B. Retire statistical significance. :3.

12. Arcaya M, Brewster M, Zigler CM, Subramanian SV. Area variations in health: A spatial multilevel modeling approach. *Health Place*. 2012;18(4):824-831. doi:10.1016/j.healthplace.2012.03.010.

13. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *The Lancet*. 2017;389(10077):1453–1463.

14. Batis C, Rivera JA, Popkin BM, Taillie LS. First-Year Evaluation of Mexico's Tax on Nonessential Energy-Dense Foods: An Observational Study. Wareham NJ, ed. *PLOS Med*. 2016;13(7):e1002057. doi:10.1371/journal.pmed.1002057.

15. Baumgartner FR, Christiani L, Epp D, Olivella S, Roach K, Shoub K. Driving while Black (and male, and young, and...): Evidence of disparities at the margin and the intersection. 2018:24.

16. Baumgartner FR, Epp DA, Shoub K. *Suspect Citizens: What 20 Million Traffic Stops Tell Us about Policing and Race*. United Kingdom ; New York, NY: Cambridge University Press; 2018.

17. Becker GS. Crime and Punishment: An Economic Approach. *J Polit Econ*. 1968:169-217.

18. Bobo LD, Thompson V. Unfair by Design: The War on Drugs, Race, and the Legitimacy of the Criminal Justice System. *Soc Res*. 2006;73(2):445-472,736,0_3.

19. Borrell LN. Editorial: Critical Race Theory: Why Should We Care about Applying It in our Research? *Ethn Dis*. 2018;28(Supp 1):215. doi:10.18865/ed.28.S1.215.

20. Braver ER. Race, Hispanic origin, and socioeconomic status in relation to motor vehicle occupant death rates and risk factors among adults. *Accid Anal Prev*. 2003;35(3):295–309.

21. Bright SB. Discrimination, death and denial: The tolerance of racial discrimination in infliction of the death penalty. *St Clara Rev*. 1994;35:433.

22. Bureau of Transportation Statistics. *Methodology for 2017 Local Area Transportation Characteristics for Households*.; 2017. https://www.bts.gov/sites/bts.dot.gov/files/docs/browse-statistical-products-and-data/surveys/224076/latch2017methodology.pdf. Accessed June 6, 2019.

23. Byrne J, Hummer D. Technology, innovation and twenty-first-century policing. https://www.researchgate.net/profile/Don_Hummer/publication/309646701_Technology_innovation_and_twenty-first-century_policing/links/581b6a2a08aea429b28fca34.pdf. Accessed December 1, 2016.

24. Campaign Zero. https://www.joincampaignzero.org/. Accessed June 11, 2019.

25. *Chicago PD Traffic Stops and Resulting Searches in 2013*. ACLU of Illinois; 2014.

26. Chief Justice Warren. *Terry v. Ohio*.(U.S. Supreme Court 1968).

27. Chin GJ. Race, The War on Drugs, and the Collateral Consequences of Criminal Conviction. *SSRN Electron J*. 2003;6:253. doi:10.2139/ssrn.390109.

28. Chomsky N. *Profit over People: Neoliberalism and Global Order*. Seven Stories Press 1st ed. New York: Seven Stories Press; 1999.

29. Colchero MA, Popkin BM, Rivera JA, Ng SW. Beverage purchases from stores in Mexico under the excise tax on sugar sweetened beverages: observational study. *BMJ*. January 2016:h6704. doi:10.1136/bmj.h6704.

30. COPS Office: Ethics and Integrity Training. http://www.cops.usdoj.gov/Default.asp?Item=2618. Accessed April 27, 2015.

31. Crime Mapping for DDACTS - Crime Mapping & Analysis NewsCrime Mapping & Analysis News. https://crimemapping.info/article/crime-mapping-ddacts/. Accessed July 17, 2018.

32. Dai Y, Shelley Brock Roth. *2017 NHTS Weighting Report*.; 2017:53.

33. Davis E. Contacts Between Police and the Public, 2015. *US Dep Justice Off Justice Programs Bur Justice Stat Spec Rep*. October 2018:33.

34. Desmond M, Papachristos AV, Kirk DS. Police Violence and Citizen Crime Reporting in the Black Community. *Am Sociol Rev*. 2016;81(5):857-876. doi:10.1177/0003122416663494.

35. Driving Safety Programs | Fayetteville, NC. https://fayettevillenc.gov/government/city-departments/police/community-policing/driving-safety-programs. Accessed May 26, 2019.

36. Driving while black: Carolina city crafts racial awareness police policy | Al Jazeera America. http://america.aljazeera.com/articles/2014/9/9/racial-profilingnorthcarolina.html. Accessed June 5, 2019.

37. Duran DG, Pérez-Stable EJ. Novel Approaches to Advance Minority Health and Health Disparities Research. *Am J Public Health*. 2019;109(S1):S8-S10. doi:10.2105/AJPH.2018.304931.

38. Engel RS, Calnon JM. Comparing Benchmark Methodologies for Police-Citizen Contacts: Traffic Stop Data Collection for the Pennsylvania State Police. *Police Q*. 2004;7(1):97-125. doi:10.1177/1098611103257686.

39. Epp CR, Maynard-Moody S, Haider-Markel DP. *Pulled over: How Police Stops Define Race and Citizenship*. Chicago ; London: The University of Chicago Press; 2014.

40. Farmer CM, Retting RA, Lund AK. Changes in motor vehicle occupant fatalities after repeal of the national maximum speed limit. *Accid Anal Prev*. 1999;31(5):537-543. doi:10.1016/S0001-4575(99)00010-X.

41. Fayetteville Police Chief Medlock announces retirement - News - The Fayetteville Observer - Fayetteville, NC. https://www.fayobserver.com/article/20160823/News/308239939. Accessed May 26, 2019.

42. *FAYETTEVILLE POLICE DEPARTMENT WRITTEN DIRECTIVES AND OPERATING PROCEDURES*. Fayetteville Police Department; 2015.

https://www.bwcscorecard.org/static/2016/policies/2015-12-18%20Fayetteville%20-%20BWC%20Policy.pdf. Accessed May 27, 2019.

43. Fayetteville ranks high in auto crashes, driving up insurance rates for drivers - News - The Fayetteville Observer - Fayetteville, NC. https://www.fayobserver.com/article/20130310/News/303109882. Accessed May 26, 2019.

44. Feldman JM, Gruskin S, Coull BA, Krieger N. Police-Related Deaths and Neighborhood Economic and Racial/Ethnic Polarization, United States, 2015–2016. *Am J Public Health*. 2019;109(3):458-464. doi:10.2105/AJPH.2018.304851.

45. Fontaine J, Leitson D, Jannetta J, Paddock E. Mistrust and Ambivalence between Residents and the Police. :22.

46. Ford CL, Airhihenbuwa CO. Commentary: Just What is Critical Race Theory and What's it Doing in a Progressive Field like Public Health? *Ethn Dis*. 2018;28(Supp 1):223. doi:10.18865/ed.28.S1.223.

47. Ford CL, Airhihenbuwa CO. Critical race theory, race equity, and public health: toward antiracism praxis. *Am J Public Health*. 2010;100(supplement 1):S30–S35.

48. Fridell LA, Police Executive Research Forum, United States, Department of Justice, Office of Community Oriented Policing Services. *Understanding Race Data from Vehicle Stops a Stakeholder's Guide*. Washington, D.C.: Police Executive Research Forum; 2005. http://www.cops.usdoj.gov/mime/open.pdf?Item=1577. Accessed November 15, 2015.

49. Galea S. An Argument for a Consequentialist Epidemiology. *Am J Epidemiol*. 2013;178(8):1185-1191. doi:10.1093/aje/kwt172.

50. Gallagher J, Fisher PJ. Deterrence in a Setting with Multiple Risks: Traffic Cameras, Vehicular Accidents, and Public Safety. 2017. http://faculty.weatherhead.case.edu/jpg75/pdfs/red_light_camera_14.pdf. Accessed July 17, 2017.

51. Gantz T, Shaver B. Traffic Safety and Communities of Color. :15.

52. García JJ, Gee GC, Jones M. A CRITICAL RACE THEORY ANALYSIS OF PUBLIC PARK FEATURES IN LATINO IMMIGRANT NEIGHBORHOODS. *Bois Rev Soc Sci Res Race*. 2016;13(02):397-411. doi:10.1017/S1742058X16000187.

53. Garrett BL, Crozier W. Driver's License Suspension in North Carolina. *SSRN Electron J*. 2019. doi:10.2139/ssrn.3355599.

54. Gee PhD Gilbert C. A Multilevel Analysis of the Relationship Between Institutional and Individual Racial Discrimination and Health Status. *Am J Public Health*. 2008;98:S48-S56.

55. Gius M. Using the synthetic control method to determine the effects of concealed carry laws on state-level murder rates. *Int Rev Law Econ*. 2019;57:1-11. doi:10.1016/j.irle.2018.10.005.

56. Golembeski C, Fullilove R. Criminal (In)Justice in the City and Its Associated Health Consequences. *Am J Public Health*. 2008;98(Suppl 1):S185-S190.

57. Goodwin A, Thomas L, Kirley B, Hall W, O'Brien N, Hill K. *Countermeasures That Work: A Highway Safety Countermeasure Guide For State Highway Safety Offices Eighth Edition*. Washington, DC: National Highway Traffic Safety Administration; 2015:567.

58. Graham MR, Kutzbach MJ, McKenzie B, others. *Design Comparison of LODES and ACS Commuting Data Products*.; 2014.

59. Green VH. *The Negro Travelers' Green Book: The Guide to Travel and Vacations*. About Comics; 1957.

60. Greenland S, Senn SJ, Rothman KJ, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol*. 2016;31(4):337-350. doi:10.1007/s10654-016-0149-3.

61. Grogger J, Ridgeway G. Testing for Racial Profiling in Traffic Stops From Behind a Veil of Darkness. *J Am Stat Assoc*. 2006;101(475):878-887. doi:10.1198/016214506000000168.

62. Gross N, Mann M. Is There a "Ferguson Effect?" Google Searches, Concern about Police Violence, and Crime in US Cities, 2014–2016. *Socius*. 2017;3:2378023117703122.

63. Hamm JA, Trinkner R, Carr JD. Fair Process, Trust, and Cooperation: Moving Toward an Integrated Framework of Police Legitimacy. *Crim Justice Behav*. 2017:0093854817710058.

64. Harper S, Charters TJ, Strumpf EC. Trends in Socioeconomic Inequalities in Motor Vehicle Accident Deaths in the United States, 1995–2010. *Am J Epidemiol*. 2015;182(7):606-614. doi:10.1093/aje/kwv099.

65. Hatzenbuehler ML, Keyes K, Hamilton A, Uddin M, Galea S. The Collateral Damage of Mass Incarceration: Risk of Psychiatric Morbidity Among Nonincarcerated Residents of High-Incarceration Neighborhoods. *Am J Public Health*. 2015;105(1):138–143.

66. Hebert PL, Sisk JE, Howell EA. When Does A Difference Become A Disparity? Conceptualizing Racial And Ethnic Disparities In Health. *Health Aff (Millwood)*. 2008;27(2):374-382. doi:10.1377/hlthaff.27.2.374.

67. Henry L, Wickham H. *Purrr: Functional Programming Tools*.; 2019. https://CRAN.R-project.org/package=purrr.

68. Hertzman C, Boyce T. How Experience Gets Under the Skin to Create Gradients in Developmental Health. *Annu Rev Public Health*. 2010;31(1):329-347. doi:10.1146/annurev.publhealth.012809.103538.

69. Hinton EK. *From the War on Poverty to the War on Crime: The Making of Mass Incarceration in America*. Cambridge, Massachusetts: Harvard University Press; 2016.

70. Hirschtick JL, Homan SM, Rauscher G, et al. Persistent and aggressive interactions with the police: potential mental health implications. *Epidemiol Psychiatr Sci*. February 2019:1-8. doi:10.1017/S2045796019000015.

71. Hughes C, Warren PY, Stewart EA, Tomaskovic-Devey D, Mears DP. Racial Threat, Intergroup Contact, and School Punishment. *J Res Crime Delinquency*. January 2017:002242781668981. doi:10.1177/0022427816689811.

72. Intervention Fact Sheets - Table of Contents | Motor Vehicle Safety | CDC Injury Center. https://www.cdc.gov/motorvehiclesafety/calculator/factsheet/index.html. Accessed July 3, 2019.

73. Jones CP. Confronting Institutionalized Racism. *Phylon 1960-*. 2002;50(1/2):7. doi:10.2307/4149999.

74. Jones CP. Invited Commentary: “Race,” Racism, and the Practice of Epidemiology. *Am J Epidemiol*. 2001;154(4):299-304. doi:10.1093/aje/154.4.299.

75. Jones CP. Toward the Science and Practice of Anti-Racism: Launching a National Campaign Against Racism. *Ethn Dis*. 2018;28(Supp 1):231. doi:10.18865/ed.28.S1.231.

76. Jones CP, Truman BI, Elam-Evans LD, et al. USING ‘‘SOCIALLY ASSIGNED IN HEALTH STATUS. *Ethn Dis*. 2008;18:496.

77. Jones NL, Breen N, Das R, Farhat T, Palmer R. Cross-Cutting Themes to Advance the Science of Minority Health and Health Disparities. *Am J Public Health*. 2019;109(S1):S21-S24. doi:10.2105/AJPH.2019.304950.

78. Jones RW, Pridemore WA. Toward an Integrated Multilevel Theory of Crime at Place: Routine Activities, Social Disorganization, and The Law of Crime Concentration. *J Quant Criminol*. November 2018. doi:10.1007/s10940-018-9397-6.

79. Julie Netherland, Helena Hansen. White opioids: Pharmaceutical race and the war on drugs that wasn’t. *BioSocieties*. 2017;12(2):239-256. doi:10.1057/biosoc.2016.6.

80. Justice Scalia. *Whren v. United States*.(U.S. Supreme Court 1996).

81. Kagawa RM, Castillo-Carniglia A, Vernick JS, et al. Repeal of comprehensive background check policies and firearm homicide and suicide. *Epidemiology*. 2018;29(4):494–502.

82. Katy Smyser. Suburban Woman Found Dead in Jail Had Previous Encounters With Police - NBC Chicago. *NBC Chicago 5*. https://www.nbcchicago.com/investigations/Suburban-Woman-Found-Dead-in-Jail-Had-Previous-Encounters-With-Police-316025661.html. Published July 16, 2015. Accessed May 27, 2019.

83. Knox D, Lowe W. The Bias Is Built In: How Administrative Records Mask Racially Biased Policing. *SSRN*. 2019:68.

84. Knuth DE. *The Art of Computer Programming*. 3rd ed. Reading, Mass: Addison-Wesley; 1997.

85. Kramer RM. TRUST AND DISTRUST IN ORGANIZATIONS: Emerging Perspectives, Enduring Questions. *Nature*. 1999;397(6715):176–180.

86. Krieger N, Chen JT, Waterman PD, Kiang MV, Feldman J. Police Killings and Police Deaths Are Public Health Data and Can Be Counted. *PLOS Med*. 2015;12(12):e1001915. doi:10.1371/journal.pmed.1001915.

87. Lee AS, Weitzer R, Martínez DE. Recent Police Killings in the United States: A Three-City Comparison. *Police Q*. 2018;21(2):196-222. doi:10.1177/1098611117744508.

88. McGregor A. Politics, Police Accountability, and Public Health: Civilian Review in Newark, New Jersey. *J Urban Health*. December 2015. doi:10.1007/s11524-015-9998-4.

89. Meares TL. The Law and Social Science of Stop and Frisk. *Annu Rev Law Soc Sci*. 2014;10(1):335-352. doi:10.1146/annurev-lawsocsci-102612-134043.

90. Meares TL, Tyler TR, Gardener J. Lawful or Fair? How Cops and Laypeople Perceive Good Policing. 105:49.

91. Methipara JC (FHWA). Mobility Challenges for Households in Poverty. 2014:4.

92. Michael Dolan Fliss. LAW ENFORCEMENT TRAFFIC STOP DISPARITY MEASUREMENT REQUIRES VEHICLE ACCESS, TRAVEL VOLUME, AND MULTI-AGENCY DRIVING CONSIDERATIONS. *J Quant Criminol*. 2019;Submission Pending.

93. Mihoko Doyle J, Kao G. Are Racial Identities of Multiracials Stable? Changing Self-Identification Among Single and Multiple Race Individuals. *Soc Psychol Q*. 2007;70(4):405-423. doi:10.1177/019027250707000409.

94. Missouri Reports Wide Racial Disparity in Traffic Stops - The New York Times. http://www.nytimes.com/2015/06/02/us/big-disparity-for-blacks-pulled-over-in-missouri.html. Accessed November 16, 2015.

95. Mohler G, Brantingham PJ, Carter J, Short MB. Reducing Bias in Estimates for the Law of Crime Concentration. *J Quant Criminol*. January 2019. doi:10.1007/s10940-019-09404-1.

96. Motor Vehicle Prioritizing Interventions and Cost Calculator for States (MV PICCS) | Motor Vehicle Safety | CDC Injury Center. https://www.cdc.gov/motorvehiclesafety/calculator/index.html. Accessed July 17, 2018.

97. Muhammad M, De Loney EH, Brooks CL, Assari S, Robinson D, Caldwell CH. "I think that's all a lie…I think It's genocide": Applying a Critical Race Praxis to Youth Perceptions of Flint Water Contamination. *Ethn Dis*. 2018;28(Supp 1):241. doi:10.18865/ed.28.S1.241.

98. NC Environmental Justice Network and NC NAACP Challenge Racially Discriminatory Swine Waste Permit – Julius L. Chambers Center for Civil Rights. https://chamberscc r.org/2019/06/25/nc-environmental-justice-network-and-nc-naacp-challenge-racially-discriminatory-swine-waste-permit/?fbclid=IwAR3v0tUasLzdDRigJT2JDptFnHh8ZkfBUzHQvQZD8ByAFmKe-Bqnhro5Rjk. Accessed June 27, 2019.

99. NC State Bureau of Investigation. Crime In North Carolina. http://crimereporting.ncsbi.gov/. Accessed June 10, 2019.

100. Nix J, Campbell BA, Byers EH, Alpert GP. A Bird's Eye View of Civilians Killed by Police in 2015: Further Evidence of Implicit Bias. *Criminol Public Policy*. 2017;16(1):309-340. doi:10.1111/1745-9133.12269.

101. Nix J, Wolfe SE. Management-level Officers' Experiences with the Ferguson Effect. *Polic Int J*. 2017. https://www.researchgate.net/profile/Justin_Nix/publication/317239395_Management-level_Officers%27_Experiences_with_the_Ferguson_Effect/links/592dd8e5a6fdcc89e7555944/Management-level-Officers-Experiences-with-the-Ferguson-Effect.pdf. Accessed July 9, 2017.

102. Nuru-Jeter A, Dominguez TP, Hammond WP, et al. "It's The Skin You're In": African-American Women Talk About Their Experiences of Racism. An Exploratory Study to Develop Measures of Racism for Birth Outcome Studies. *Matern Child Health J*. 2009;13(1):29-39. doi:10.1007/s10995-008-0357-x.

103. Okun T. *The Emperor Has No Clothes: Teaching about Race and Racism to People Who Don't Want to Know*. Charlotte, N.C: Information Age Pub; 2010.

104. Open Data Policing. https://opendatapolicing.com/nc/. Accessed June 11, 2019.

105. Pebesma E. Simple Features for R: Standardized Support for Spatial Vector Data. *R J*. 2018. https://journal.r-project.org/archive/2018/RJ-2018-009/index.html.

106. Peralta E, Corley C. The driving life and death of Philando Castile. *NPR Morning Ed*. 2016.

107. Prieto Curiel R, González Ramírez H, Bishop SR. A novel rare event approach to measure the randomness and concentration of road accidents. Deng Y, ed. *PLOS ONE*. 2018;13(8):e0201890. doi:10.1371/journal.pone.0201890.

108. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2018. https://www.R-project.org/.

109. Rehkopf DH, Basu S. A New Tool for Case Studies in Epidemiology—the Synthetic Control Method: *Epidemiology*. 2018;29(4):503-505. doi:10.1097/EDE.0000000000000837.

110. RICHARD DELGADO JS. *CRITICAL RACE THEORY*. Place of publication not identified: DEV Publishers & DISTRIBU; 2016.

111. Richard Stradling. NC driver's licenses leave a space for race at request of Native Americans | Raleigh News & Observer. *The News & Observer*. https://www.newsobserver.com/news/traffic/article198464504.html. Published February 6, 2018. Accessed June 6, 2019.

112. Ridgeway G. Assessing the effect of race bias in post-traffic stop outcomes using propensity scores. *J Quant Criminol*. 2006;22(1):1–29.

113. Ritchey M, Nicholson-Crotty S. Deterrence Theory and the Implementation of Speed Limits in the American States: Ritchey/Nicholson-Crotty: Deterrence Theory and Speed Limits. *Policy Stud J*. 2011;39(2):329-346. doi:10.1111/j.1541-0072.2011.00410.x.

114. Robbins M, Davenport S. *Microsynth: Synthetic Control Methods with Micro- And Meso-Level Data*.; 2018. https://CRAN.R-project.org/package=microsynth.

115. Rodriguez D, Kunard L, Johnson W, LaRochelle J, Thorkildsen Z. Assessment Report on the Fayetteville Police Department. :132.

116. Rojek J, Rosenfeld R, Decker S. The Influence Of Driver'S Race on Traffic Stops in Missouri. *Police Q*. 2004;7(1):126-147. doi:10.1177/1098611103260695.

117. Rose G. Sick individuals and sick populations. *Int J Epidemiol*. 2001;30(3):427–432.

118. Rothstein R. *The Color of Law: A Forgotten History of How Our Government Segregated America*. Liveright Publishing; 2017.

119. Sarah A. Seo. The New Public. *Yale Law J*. 2016;125:1616-1671.

120. Sharon LaFraniere, Andrew W. Lehren. The Disproportionate Risks of Driving While Black - The New York Times. *The New York Times*. http://www.nytimes.com/2015/10/25/us/racial-disparity-traffic-stops-driving-black.html?_r=0. Published October 24, 2015. Accessed November 16, 2015.

121. Shjarback JA, Pyrooz DC, Wolfe SE, Decker SH. De-policing and crime in the wake of Ferguson: Racialized changes in the quantity and quality of policing among Missouri police departments. *J Crim Justice*. 2017;50:42-52. doi:10.1016/j.jcrimjus.2017.04.003.

122. Smith BW, Holmes MD. Police use of excessive force in minority communities: A test of the minority threat, place, and community accountability hypotheses. *Soc Probl*. 2014;61(1):83–104.

123. Stang A, Poole C, Kuss O. The ongoing tyranny of statistical significance testing in biomedical research. *Eur J Epidemiol*. 2010;25(4):225-230. doi:10.1007/s10654-010-9440-x.

124. Subramanian S, Jones K, Kaddour A, Krieger N. Response: The value of a historically informed multilevel analysis of Robinson's data. *Int J Epidemiol*. 2009;38(2):370-373. doi:10.1093/ije/dyn354.

125. Sudbury J, Okazawa-Rey M, eds. *Activist Scholarship: Antiracism, Feminism, and Social Change*. Boulder: Paradigm Publishers; 2009.

126. Tal G, Handy S. The Travel Behavior of Immigrants and Race/Ethnicity Groups: An Analysis of the 2001 National Household Transportation Survey. :39.

127. Texas troopers ticketing Hispanic drivers as white. https://www.kxan.com/news/investigations/texas-troopers-ticketing-hispanics-motorists-as-white/1156475533. Accessed May 27, 2019.

128. The People's Institute for Survival and Beyond – Undoing Racism. https://www.pisab.org/. Accessed July 14, 2019.

129. Top two Fayetteville police officials leave amid controversy. https://www.ajc.com/news/top-two-fayetteville-police-officials-leave-amid-controversy/XxV07BLN8xzLXuXn1hHE2H/. Accessed May 26, 2019.

130. Tsai AC, Venkataramani AS. Syndemics and Health Disparities: A Methodological Note. *AIDS Behav*. 2016;20(2):423-430. doi:10.1007/s10461-015-1260-2.

131. UNC Highway Safety Research Center. North Carolina Crash Data. http://nccrashdata.hsrc.unc.edu/datatool.cfm. Accessed June 10, 2019.

132. US Department of Justice, Civil Rights Division. *Investigation of the Ferguson Police Department*.; 2015.

133. US Department of Transportation: Federal Highway Administration National Household Travel Survey. http://nhts.ornl.gov/index.shtml. Accessed March 11, 2016.

134. Vandenbroucke JP, Poole C, Schlesselman JJ, Egger M. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. *PLoS Med*. 2007;4(10):27.

135. VanderWeele TJ, Robinson WR. On the Causal Interpretation of Race in Regressions Adjusting for Confounding and Mediating Variables: *Epidemiology*. 2014;25(4):473-484. doi:10.1097/EDE.0000000000000105.

136. Vaughan D, Dancho M. *Furrr: Apply Mapping Functions in Parallel Using Futures*.; 2019. https://github.com/DavisVaughan/furrr.

137. Walker K. *Tidycensus: Load US Census Boundary and Attribute Data as "tidyverse" and 'Sf'-Ready Data Frames*.; 2019. https://CRAN.R-project.org/package=tidycensus.

138. Ward JB, Gartner DR, Keyes KM, Fliss MD, McClure ES, Robinson WR. How do we assess a racial disparity in health? Distribution, interaction, and interpretation in epidemiological studies. *Ann Epidemiol*. 2019;29:1-7. doi:10.1016/j.annepidem.2018.09.007.

139. We Crunched the Numbers on Race and Traffic Stops in the County Where Sandra Bland Died | Mother Jones. http://www.motherjones.com/politics/2015/07/traffic-stops-black-people-waller-county. Accessed November 15, 2015.

140. Wickham H. A layered grammar of graphics. *J Comput Graph Stat*. 2010;19(1):3–28. doi:10.1198/jcgs.2009.07098.

141. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York; 2016. http://ggplot2.org.

142. Wickham H. Rvest: Easily harvest (scrape) web pages. *R Package Version 03*. 2015;1.

143. Wickham H. *Tidyverse: Easily Install and Load the "Tidyverse."*; 2017. https://CRAN.R-project.org/package=tidyverse.

144. Wilkinson L, Wills G. *The Grammar of Graphics*. 2nd ed. New York: Springer; 2005.

145. Withrow BL, Williams H. Proposing a Benchmark Based on Vehicle Collision Data in Racial Profiling Research. *Crim Justice Rev*. August 2015. doi:10.1177/0734016815591819.

146. Xu Y. Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. *Polit Anal*. 2017;25(01). doi:10.1017/pan.2016.2.

147. Zimmerman FJ. Habit, custom, and power: A multi-level theory of population health. *Soc Sci Med*. 2013;80:47-56. doi:10.1016/j.socscimed.2012.12.029.

148. Zingraff MT, Mason M, Smith W, et al. Evaluating North Carolina State Highway Patrol Data: Citation. *Warn Searches In*. 1998.

149. Zuberi T, Bonilla-Silva E, eds. *White Logic, White Methods: Racism and Methodology*. Lanham: Rowman & Littlefield Publishers; 2008.