Online Appendix for

Fines, Fees, and Disparities: A Link Between Municipal Reliance on Fines and Racial Disparities in Policing

Table of Contents

Appendix A: List of North Carolina Municipal Police Departments

There are 290 agencies that report individual level traffic stop information to the North Carolina State Bureau of Investigations of which there are 163 municipal police departments. Excluded agencies include state agencies, county sheriff's offices, and university and hospital police among others. Of the 163 municipal police departments, 161 have a corresponding match in the data set of municipalities that report their budget information to the state treasurer's office. 126 met our criteria for inclusion, while 35 did not. These are listed below.

- **Matched Municipal Police Departments Meeting Criteria (126 Departments)**

  Aberdeen, Albemarle, Andrews, Apex, Archdale, Asheboro, Asheville, Atlantic Beach, Biltmore Forest, Blowing Rock, Boone, Burgaw, Burlington, Candor, Carrboro, Cary, Chapel Hill, Charlotte-Mecklenburg, Claremont, Clayton, Cleveland, Concord, Conover, Cornelius, Davidson, Duck, Dunn, Durham, Eden, Elizabeth City, Fayetteville, Fletcher, Fuquay-Varina, Garner, Gastonia, Goldsboro, Graham, Greensboro, Greenville, Havelock, Henderson, Hendersonville, Hickory, High Point, Hillsborough, Holly Ridge, Holly Springs, Hope Mills, Huntersville, Jacksonville, Kannapolis, Kenly, Kernersville, Kings Mountain, Kinston, Kitty Hawk, Knightdale, Kure Beach, Lake Lure, Leland, Lenoir, Lexington, Lincolnton, Lumberton, Madison, Maggie Valley, Manteo, Mars Hill, Matthews, Mayodan, Mebane, Mint Hill, Monroe, Mooresville, Morganton, Morrisville, Mount Airy, Mount Gilead, Mount Holly, Nags Head, New Bern, Newton, Norlina, North Topsail Beach, North Wilkesboro, Ocean Isle Beach, Parkton, Pine Knoll Shores, Pinehurst, Pineville, Pittsboro, Raleigh, Reidsville, Richlands, Roanoke Rapids, Robersonville, Rocky Mount, Rolesville, Saint Pauls, Salisbury, Sanford, Shallotte, Shelby, Smithfield, Southern Pines, Spring Lake, Stallings, Statesville, Surf City, Sylva, Tarboro, Taylorsville, Thomasville, Topsail Beach, Troutman, Wake Forest, Washington, Waxhaw, White Lake, Wilkesboro, Wilmington, Wilson, Winston-Salem, Wrightsville Beach, Youngsville, Zebulon

- **Matched Municipal Police Departments Not Meeting Criteria (35 Departments)**

  Bailey, Bald Head Island, Banner Elk, Beech Mountain, Belmont, Bunn, Columbus, Enfield, Highlands, Indian Beach, Kenansville, Laurinburg, Littleton, Middlesex, Murfreesboro, Murphy, Newland, Old Fort, Pembroke, Pilot Mountain, Pinetops, Red Springs, Rowland, Saluda, Seven Devils, Sharpsburg, Spruce Pine, Stantonsburg, Star, Stoneville, Sugar Mountain, Sunset Beach, Tryon, Warrenton, Weldon

- **Not Matched Municipal Police Departments (2 Departments)**

  Gaston County, Village of Misenheimer

Appendix B: Finding Contraband and Seizing Property During a Traffic Stop

In the body of the paper, we focus on whether a driver is searched and if searched whether contraband is found. However, there are two potential concerns with this analysis that are addressed in this appendix. First, we do not directly test whether any property or assets are seized by police officers during a stop. In other words, we are not testing whether any property or assets are *forfeited* by a member of the public. Second, we may not be fully capturing whether the burdens are disproportionally born by drivers of different races as suggested by the hypothesis. To assuage these concerns, we test the robustness of the results of the contraband analysis – and test of the third hypothesis in the paper – in two ways. To preview these results, the same patterns are detected for who is found with contraband and who has property and assets seized: the results hold.

First, we replicate the contraband analysis presented in the second model in Table 1 in the paper by predicting whether any property is seized as a part of the traffic stop *if a search is conducted.* This allows us to more directly test the specific concept of increased reliance leading to increased forfeitures. Here we test essentially the same hypothesis as with the contraband analysis: (Hypothesis 3a) as reliance on fees increases and searches of whites decline, we should observe an increased probability of property being seized following a search of a white driver but not of a black driver. The original contraband analysis is shown in the Table B1 as model 2, while the updated seizure analysis is shown as model 4. Across both models, we observe the same patterns and support for hypotheses 3 and 3a.

Second, we re-estimate the contraband and seizure regressions such that they are not conditional on a search occurring. In this case, if white drivers see disproportionally lighter outcomes (i.e., black drivers experience harsher outcomes) when there is increased revenue reliance, then we should observe the sign associated with a driver being black and the interaction terms between driver race and revenue reliance to reverse. This is because if black drivers are more likely to be subjected to the actions that lead to contraband being found and property being siezed – i.e., they are search at higher rates – then if one does not account for that step we should see black drivers to be more likely to be found with contraband and have their property seized. This leads to the following hypothesis: (*Hypothesis 4*) if white drivers are less likely to be searched than black drivers, then black drivers should be more likely to be found with contraband and have their property seized, when looking at the non-conditional relationship. Once again, this is observed.

In all of the models discussed above and presented in Table B1, we use data from the North Carolina Traffic stops data set described in the paper. In each, the unit of analysis is an individual stop, and we include the same variables as described in the paper that are associated with the models presented in Table 1. Finally, logistic regressions are estimated.

Table B1. Explaining Who is Found with Contraband and Who has Property Seized

	(1) Contraband	(2) Contraband | Search	(3) Seizure	(4) Seizure | Search
Intercept	-6.12**	0.56	-6.84**	-0.90
	(0.38)	(0.56)	(0.58)	(0.66)
Black Driver	0.53**	-0.02	0.44**	-0.15**
	(0.02)	(0.02)	(0.02)	(0.02)
Prop. of Total Revenue	-4.18**	8.14**	-6.64**	3.98**
	(0.88)	(1.13)	(1.16)	(1.30)
Dif. in Total Revenue	0.15**	0.05	0.19**	0.03
	(0.02)	(0.03)	(0.03)	(0.04)
Revenue & Expenditures Dif.	0.25**	0.14**	0.31**	0.14**
	(0.04)	(0.05)	(0.05)	(0.05)
Black * Prop. of Total Rev.	6.16**	-6.45**	7.02**	-3.21*
	(0.90)	(1.16)	(1.19)	(1.34)
Black * Dif. in Total Rev.	0.12**	0.04	0.09*	0.05
	(0.03)	(0.04)	(0.04)	(0.05)
Black * Dif. Rev. & Expend.	0.16**	-0.10	0.08	-0.08
	(0.04)	(0.06)	(0.06)	(0.07)
Male Driver	1.04**	0.02	1.00**	0.01
	(0.01)	(0.01)	(0.01)	(0.02)
Driver Age	-0.05**	-0.01**	-0.05**	-0.01**
	(<0.01)	(<0.01)	(<0.01)	(0.00)
Latinx Drivers & Interactions	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes
Municipal Fixed Effects	Yes	Yes	Yes	Yes
AIC	552,761	209,106	335,158	157,690
BIC	555,036	210,816	337432	159,390
Log Likelihood	-276,210	-104,382	-167408	-78,675
Num. obs.	4,409,385	162,752	4,409385	162,750

Note: ** indicates $p < 0.01$ and * indicates $p < 0.05$. Entries are coefficients from a logistic regression, with standard errors in parentheses below. FFF stands for "fines, fees, and forfeitures." "Other controls" are: initial stop purpose month of stop, day of week of the stop, year of stop, and hour of stop. Observations that are extreme outliers with respect to revenue measures are excluded.

Appendix C: Generating the Time Windows for the Traffic Stops Data

Each observation in the data must meet the thresholds laid out in the body of the paper: 10,000 stops, including 100 white drivers and 100 black drivers must also be stopped. Data on traffic stops comes from two different types of data sources: (1) data sets from states that include information about *individual* stops or micro-level data sets and (2) reports and data sets from individual agencies (ex. the Austin police department in Texas) and for states (ex. Missouri) that simply report aggregate values or macro-level data sets. For the macro-level data sets, we simply institute thresholds for inclusion. For the micro-level data sets, we institute an aggregation process to maximize the number of possible observations in the study. This distinction between the data sets is made in order to better understand what affect this aggregation may or may not have on associated variables *and* to ensure that there are enough observations to validate the measure of search disparities used in the paper. This test is discussed in Appendix B. Here is a longer discussion of the aggregation process for agencies where we have micro-level data.

If the thresholds were not met for agencies in states where we have individual stop data, we added the following year for the same agency until the threshold was met. For example, if a given agency did not have more than 10,000 total stops and over 100 stops for each race category in its first year of data (say, 2005), we would add data from the next year, in this case 2006. If this combination met the thresholds, it would constitute its own observation in the macro level dataset, and the process would begin again with 2007. If the 2005 and 2006 combination did not break the threshold then we would combine 2005 and 2006 observation with the 2007 observation, repeating this process until the threshold was met (if the threshold was not met for the combination in the last year data was available for that agency, then the observation was dropped).

Table C1 reports the number of agency-year observations that initially met the thresholds, as well as the number of observations derived from the method described above. As the table makes clear, we increased the number of usable individual stops from 29,027,595 to 44,654,524.

Table C1: Summary of observations for different aggregation methods

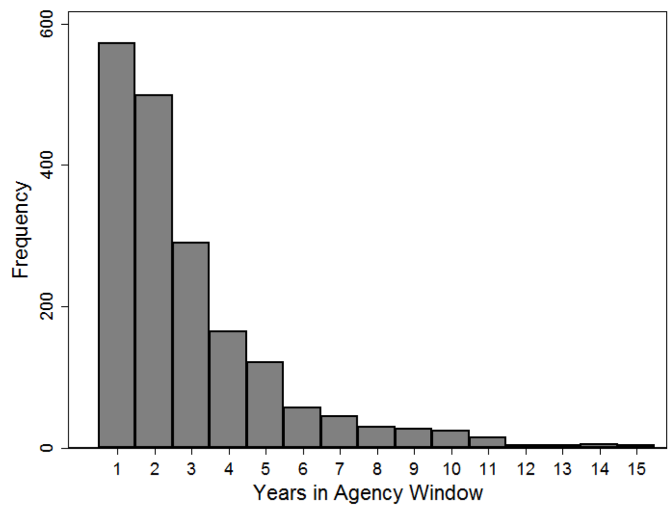
	Agency-Year Observations		Agency-Window
State	All Obs.	Obs. Above Thresholds	Obs. Above Thresholds
IL	25,767,976	11,605,503	22,404,446
MD	2,854,963	2,193,528	2,548,013
CT	859,923	349,088	461,517
NC	20,235,751	14,879,476	19,240,548
Total	49,718,613	29,027,595	44,654,524

This process did create observations with different time boundaries. Table C2 reports the summary statistics for the time frame (calculated as the end year for the observation minus the start year for the observation plus 1) for the macro level observations. The time range for the observations ranging from one year up to 15 years, with an average of 2.93 years. Figure C1 plots the distribution of the number of years that each observation includes in the agency-window.

Table C2: Summary Statistics for the Number of Years Each Observation Spans

	Observations	Mean	Std. Dev.	Minimum	Maximum
Time Frame	1857	2.930	2.346	1	15

Figure C1: Frequency Distribution of the Number of Years Each Observation Spans



Appendix D: Comparison of Black:White Search Rate Ratios and Logistic Odds Ratios for Black Drivers

Many law enforcement agencies across the country make some basic traffic stop data available, but four states mandate the collection and public availability of detailed contextual information about each stop from every police agency, not only the highway patrol. These states make available not only annual summaries, but the full micro-level databases, with a record for each individual traffic stop. The question is what the appropriate way to summarize potential disparities in treatment of black and white drivers using the available data.

Here we show that a simple search rate ratio captures the same information as approaches that account for other explanations as to why a driver is searched following a traffic stop. For the simple SRR to be said to capture the same information as a more robust specification, the two measures should be highly correlated and statistically linked. If the more robust specification sees the effect of race sufficiently attenuated by other covariates, then the two will not be sufficiently correlated nor can one be used to predict the other. To test this, we use the micro-level data associated with the agency-time windows described in Appendix C.

Using that data, we fit separate logistic regressions for each agency-window where we have micro-level data. The regressions explain whether a driver is searched using the most information available for a given time window and agency. However, because different states require the collection of different pieces of information about the driver and stop, the variables used in the regressions for agencies within different states are slightly different. Table D1 summarizes the data used. Table D2 summarizes the variables included in the regressions based on the state. Figure D1 presents histograms of the distribution of each of the measures. Note that they look very similar: high peak, long right tail, almost all values greater than 1 or equality.

Table D1. Traffic stops, searches, and search rates by state and race

					Percent Searched		
State	Years	Agency-Periods	Stops	Searches	Total	White	Black
CT	2013-15	15	461,517	14,796	3.2%	2.0%	7.3%
IL	2004-14	1,222	21,958,971	924,674	4.2%	2.5%	7.3%
MD	2013-16	68	2,548,013	75,071	3.2%	2.7%	4.1%
NC	2002-16	552	19,240,543	517,621	5.4%	3.9%	7.3%
Total		1,857	44,209,044	1,532,162			

Table D2. Summary of variables available by state

Variable	CT	IL	MD	NC
Race-Ethnicity	X	X	X	X
Gender	X	X	X	X
Driver Age	X	X	X	X
Stop Purpose	X	X	X	X
Hour of Day	X	X	X	X
Day of Week	X	X	X	X
Out of State	X			
High Disparity Officer	X		X	X
Vehicle Age		X		

Note: X indicates the variable was included. A blank indicates the variable was not available.

Figure D1. Search Rate Ratios Distributions

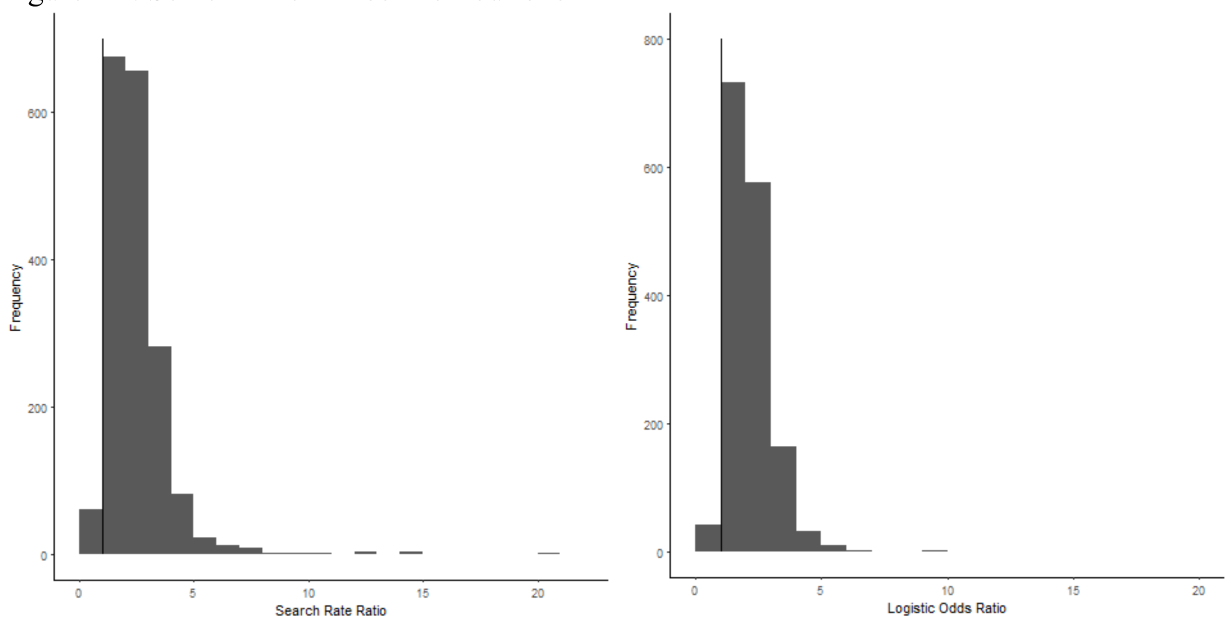


Table D3 summarizes the two validation tests run where the SRRs and odds ratios from the logistic regressions associated with the driver being black are compared. The two tests are: (1) calculating the correlation between the scores (presented in the second column), and (2) regressing the SRR on the odds ratio to see if the associate coefficient is statistically significant (presented in the third and fourth columns). As can be seen, the two measures are almost perfectly correlated and predict each other with fairly high accuracy. However, the coefficient mapping the SRRs onto the odds ratios are not 1.00, which would indicate a direct translation. Instead each is between 0.85 and 0.89, which indicates that there is some value in the extra information. Despite this, the relative increase or decrease in disparity is picked up in the simpler measure. In turn, this means that the easily computed and widely available SSR is a robust measure of racial disparity.

Table D3. Comparing Search-Rate Ratios and Odds Ratios from Logistic Regressions

State	Pearson Correlation	OLS Coef. (St. Err.)	Adjusted R^2	N
All	0.98	0.86 (0.00)	0.97	1,557
IL	0.98	0.85 (0.00)	0.97	1,002
MD	0.98	0.85 (0.01)	0.96	59
NC	0.99	0.89 (0.01)	0.99	482

Note: All OLS coefficients are statistically significant at the 0.05 level. Separate results are not presented for Connecticut because there are only five observations that meet our thresholds in that state.

Appendix E: Inclusion Thresholds for Cross-State Analysis

In the analysis presented in the paper, we include all observations where we have more than one year of data for a municipality and where the current and previous search rate is above zero. However, we also test our hypothesis using a more conservative dataset where we exclude outliers (SRR, proportion of revenue, and difference in revenue) and restrict the data to only observations where the SRR is calculated using one year of data. In both of these cases, we present the results in the body of the paper see the same result. However, we do not show the data summary for both. Here we do that. The first sub-table summarizes the entire data set, the second summarizes the slightly constricted data set used in the main analysis in the paper, and the third summarizes the very conservative data asset used in associated robustness check.

Table E1. Overview of the Aggregate Data Set by Inclusion/Exclusion Criteria

a. All Observations where a Match Between the Two Underlying Data Sets Exists

State	N	Municipalities	Stops	Mean SRR	Mean Difference in Revenue per Capita	Mean Revenue Proportion_{t-1}
IL	151	34	3,597,961	3.22	$0.06	0.0178
NC	68	14	2,009,142	2.02	$0.03	0.0000
OR	7	1	391,850	2.52	$0.10	0.0001
TX	3	1	561,730	3.17	$0.41	0.0094
Overall	229	50	6,560,413	2.84	$0.05	0.0118

b. Observations where Outliers are Included and Municipalities Appear Multiple Times

State	N	Municipalities	Stops	Mean SRR	Mean Difference in Revenue per Capita	Mean Revenue Proportion_{t-1}
IL	143	26	3,483,283	3.26	$0.06	0.0179
NC	65	11	1,960,441	2.00	$0.02	0.0000
OR	7	1	391,850	2.52	$0.10	0.0001
TX	3	1	561,730	3.17	$0.41	0.0094
Overall	218	39	6,397,304	2.86	$0.06	0.0118

c. Observations where Outliers are Included and Municipalities Appear Multiple Times

State	N	Municipalities	Stops	Mean SRR	Mean Difference in Revenue per Capita	Mean Revenue Proportion_{t-1}
IL	82	18	1,376,827	2.97	$0.06	0.0102
NC	59	10	1,829,137	2.03	$0.03	0.0000
OR	7	1	391,850	2.52	$0.10	0.0001
TX	2	1	328,882	3.32	-$0.11	0.0079
Overall	150	30	3,926,696	2.58	$0.05	0.0057