

Suspect Citizens

What 20 Million Traffic Stops Tells Us about Policing and Race

Frank R. Baumgartner, Derek A. Epp, and Kelsey Shoub
Cambridge University Press, 2018

Appendix to Chapter 8

Black Political Power and Disparities in Policing

Summary Statistics

Table 8-A1. Summary Statistics for Variables Included in Regression in Table 8.2

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	N
Investigatory Stops Ratio	0.83	1.10	1.18	1.20	1.28	1.71	508
Political Power	-0.82	-0.51	0.03	0.20	0.73	2.74	508
Log(Population)	7.15	9.79	10.66	10.69	11.54	13.51	508
Crime per Hundred	1.86	7.72	10.79	10.84	13.41	36.10	508
% Below Poverty	0.03	0.15	0.18	0.17	0.21	0.30	508

Political Power and Additional Variables

We looked at two additional variables that may influence disparities but did not include them in the analysis in this chapter because of missing data problems. These are the proportion of the police force that is black and the level of residential segregation. Information on the proportion of the police force can be drawn from the Law Enforcement Management and Administrative Statistics (LEMAS) datasets, which is administered every few years by the US Department of Justice. Information on the level of segregation in a city or town can be drawn from the diversity and disparities datasets provided by Brown University, based on US Census data (<https://s4.ad.brown.edu/projects/diversity/Data/data.htm>). However, each variable is missing for a number of cities. This missingness is not at random but rather linked to the level of political

power of the black community of the city and the population of the city, as shown in Table 8-A2. As a result, we exclude these two variables from the analysis presented in the chapter.

Table 8-A2. Changing Statistics Based on Changing Datasets

	N	Cities	Political Power	Log(Population)	Invest. Stops	Ratio
Baseline	508	86	0.06	10.11		1.18
Baseline + LEMAS	257	38	0.20	10.69		1.20
Baseline + Segregation	267	26	0.39	10.63		1.19
All	174	25	0.43	11.02		1.21

Looking at the 499 cases included in our analysis, we see for example that the political power variable has a mean value of 0.06. As we add the LEMAS data on the composition of the police force, we lose almost half of the observations (N = 258, just 38 cities out of 86), and when we add both those variables and the census data on segregation, we are down to just 25 cities, 174 observations. These are not randomly selected, however, but move us to 0.42 on the political power variable, and almost an entire point on the log of the population. We look forward to investigating the impact of these variables in another project but cannot have confidence in the analysis for a single state as we lose too many observations, and these are not randomly drawn but highly associated with population size and our key variable of black political power.

Amount of Variation Explained by Agency & Year for Regressions

Table 8-A3 presents the proportion of variation explained by agency and year respectively. In each case, we can see that there is little variation explained by over-time change. Additionally, we can see that much of the variation is explained by agency.

Table 8-A3. Analysis of Variation by Agency and Year for Regressions in Table 8.2.

	Agency	Year	Residual
Search Ratio	0.44	0.01	0.55
Light Outcome Ratio	0.47	0.02	0.51
Citation Ratio	0.56	0.02	0.42
Arrest Ratio	0.27	0.01	0.72

Assessment of the Racial Threat Hypothesis

To test for whether inclusion of a variable explicitly measuring the extent of racial threat influences the estimated relationship, we fit the regressions presented in the chapter with one additional variable: extent of racial threat. Then the fit statistics are compared. In each model, racial threat is calculated as:

$$Racial\ Threat = |Black\ \% \ of\ the\ Population - Threshold|$$

We use four different thresholds. These are: 20, 30, 40, and 50. Table 8-A4 presents the results for the search rate ratio variable. Table 8-A5 presents the results for the light outcome rate ratio variable. Table 8-A6 presents the results for the citation rate ratio variable. Table 8-A7 presents the results for the arrest rate ratio variable.

Table 8-A4. Robustness of the Search Rate Ratio Regression Given the Racial Threat Hypothesis

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	-0.202 (0.683)	-0.354 (0.704)	-0.628 (0.726)	-0.744 (0.773)	-1.095 (0.812)
Political Power	-0.240* (0.106)	-0.305* (0.129)	-0.244* (0.105)	-0.154 (0.121)	-0.016 (0.155)
Investigatory Stops Ratio	0.900** (0.328)	0.867** (0.330)	0.850** (0.329)	0.885** (0.328)	0.911** (0.327)
Log(Population)	0.105 (0.061)	0.116 (0.062)	0.133* (0.063)	0.131* (0.063)	0.130* (0.061)
Crime per 100	0.000 (0.014)	-0.001 (0.014)	0.002 (0.014)	0.003 (0.014)	0.006 (0.014)
% Below Poverty	0.104 (1.048)	0.118 (1.047)	0.119 (1.042)	0.153 (1.046)	0.244 (1.042)
Threat, 20		0.008 (0.009)			
Threat, 30			0.013 (0.008)		
Threat, 40				0.013 (0.009)	
Threat, 50					0.019* (0.010)
AIC	1362.315	1371.169	1369.444	1369.8	1367.805
BIC	1396.159	1409.244	1407.518	1407.875	1405.88
Log Likelihood	-673.157	-676.585	-675.722	-675.9	-674.903
Num. obs.	508	508	508	508	508
Num. groups: City	86	86	86	86	86
Var: City (Intercept)	0.313	0.312	0.308	0.31	0.305
Var: Residual	0.654	0.655	0.653	0.653	0.652

Note: Entries are regression coefficients, with standard errors in parentheses. ** prob. < 0.01 * prob. < 0.05.

Table 8-A4 shows that the model excluding the racial threat variable performs the best when predicting search rate ratios: the AIC and BIC are minimized in this regression, and the log likelihood is maximized. Additionally, regardless of specification of a racial threat threshold, the investigatory stop rate ratio variable is consistently statistically significant at the 0.05 level, and the coefficient moves a negligible amount. The political power variable consistently points in the correct direction, but is inconsistently statistically significant.

Table 8-A5. Robustness of the Light Outcome Rate Ratio Regression Given the Racial Threat Hypothesis

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.673** (0.085)	0.706** (0.087)	0.727** (0.090)	0.741** (0.096)	0.730** (0.102)
Political Power	-0.041** (0.013)	-0.027 (0.016)	-0.041** (0.013)	-0.052** (0.015)	-0.056** (0.020)
Investigatory Stops Ratio	0.147** (0.042)	0.154** (0.042)	0.154** (0.042)	0.149** (0.042)	0.146** (0.042)
Log(Population)	0.015* (0.007)	0.012 (0.008)	0.011 (0.008)	0.011 (0.008)	0.013 (0.008)
Crime per 100	-0.003 (0.002)	-0.002 (0.002)	-0.003 (0.002)	-0.003 (0.002)	-0.003 (0.002)
% Below Poverty	0.278* (0.130)	0.275* (0.130)	0.277* (0.129)	0.272* (0.130)	0.270* (0.130)
Threat, 20		-0.002 (0.001)			
Threat, 30			-0.002 (0.001)		
Threat, 40				-0.002 (0.001)	
Threat, 50					-0.001 (0.001)
AIC	-701.915	-690.346	-690.758	-690.314	-689.376
BIC	-668.071	-652.272	-652.684	-652.239	-651.301
Log Likelihood	358.958	354.173	354.379	354.157	353.688
Num. obs.	508	508	508	508	508
Num. groups: City	86	86	86	86	86
Var: City (Intercept)	0.005	0.005	0.005	0.005	0.005
Var: Residual	0.011	0.011	0.011	0.011	0.011

Note: Entries are regression coefficients, with standard errors in parentheses. ** prob. < 0.01 * prob. < 0.05.

Table 8-A5 shows that the model excluding the racial threat variable performs the best when predicting search rate ratios: the AIC and BIC are minimized in this regression, and the log likelihood is maximized. Additionally, regardless of specification of a racial threat threshold, the investigatory stop rate ratio variable is consistently statistically significant at the 0.05 level, and the coefficient moves a negligible amount. The political power variable consistently points in the correct direction, and is typically statistically significant.

Table 8-A6. Robustness of the Citation Rate Ratio Regression Given the Racial Threat Hypothesis

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	1.310** (0.068)	1.314** (0.070)	1.326** (0.073)	1.320** (0.077)	1.300** (0.082)
Political Power	0.034** (0.010)	0.036** (0.013)	0.034** (0.010)	0.033** (0.012)	0.036* (0.015)
Investigatory Stops Ratio	-0.131** (0.031)	-0.130** (0.031)	-0.129** (0.031)	-0.131** (0.031)	-0.131** (0.031)
Log(Population)	-0.015* (0.006)	-0.015* (0.006)	-0.016* (0.006)	-0.015* (0.006)	-0.014* (0.006)
Crime per 100	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)
% Below Poverty	-0.131 (0.104)	-0.131 (0.105)	-0.131 (0.104)	-0.132 (0.105)	-0.129 (0.105)
Threat, 20		-0.000 (0.001)			
Threat, 30			-0.000 (0.001)		
Threat, 40				-0.000 (0.001)	
Threat, 50					0.000 (0.001)
AIC	-1016.019	-1001.845	-1001.955	-1001.809	-1001.99
BIC	-982.175	-963.771	-963.881	-963.735	-963.916
Log Likelihood	516.009	509.923	509.978	509.905	509.995
Num. obs.	508	508	508	508	508
Num. groups: City	86	86	86	86	86
Var: City (Intercept)	0.003	0.003	0.003	0.003	0.003
Var: Residual	0.006	0.006	0.006	0.006	0.006

Note: Entries are regression coefficients, with standard errors in parentheses. ** prob. < 0.01 * prob. < 0.05.

Table 8-A6 shows that the model excluding the racial threat variable performs the best when predicting search rate ratios: the AIC and BIC are minimized in this regression, and the log likelihood is maximized. Additionally, regardless of specification of a racial threat threshold, the investigatory stop rate ratio variable is consistently statistically significant at the 0.05 level, and the coefficient moves a negligible amount. Finally, regardless of specification of a racial threat threshold, the political power variable is consistently statistically significant at the 0.05 level.

Table 8-A7. Robustness of the Arrest Rate Ratio Regression Given the Racial Threat Hypothesis

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	1.251*	1.358*	1.039	0.966	0.594
	(0.521)	(0.541)	(0.555)	(0.593)	(0.632)
Political Power	-0.022	0.024	-0.024	0.031	0.170
	(0.081)	(0.102)	(0.081)	(0.094)	(0.126)
Investigatory Stops Ratio	0.470	0.490	0.425	0.443	0.448
	(0.281)	(0.284)	(0.283)	(0.282)	(0.282)
Log(Population)	-0.021	-0.028	-0.005	-0.006	-0.002
	(0.045)	(0.047)	(0.047)	(0.048)	(0.047)
Crime per 100	-0.005	-0.004	-0.004	-0.003	-0.001
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
% Below Poverty	1.107	1.083	1.118	1.129	1.189
	(0.792)	(0.799)	(0.796)	(0.802)	(0.807)
Threat, 20		-0.005			
		(0.007)			
Threat, 30			0.007		
			(0.006)		
Threat, 40				0.007	
				(0.007)	
Threat, 50					0.016*
					(0.008)
AIC	1244.475	1254.091	1253.516	1253.475	1250.534
BIC	1278.318	1292.166	1291.59	1291.55	1288.608
Log Likelihood	-614.237	-618.046	-617.758	-617.738	-616.267
Num. obs.	508	508	508	508	508
Num. groups: City	86	86	86	86	86
Var: City (Intercept)	0.133	0.137	0.135	0.14	0.144
Var: Residual	0.555	0.554	0.554	0.552	0.548

Note: Entries are regression coefficients, with standard errors in parentheses. ** prob. < 0.01 * prob. < 0.05.

Table 8-A7 shows that the model excluding the racial threat variable performs the best when predicting search rate ratios: the AIC and BIC are minimized in this regression, and the log likelihood is maximized.