

Why a Small Handful of Counties Generates the Bulk of US Death Sentences A Theory of Self-Reinforcement with Three Statistical Tests

Frank R. Baumgartner, University of North Carolina at Chapel Hill
Janet M. Box-Steffensmeier, Ohio State University
Benjamin W. Campbell, Ohio State University
Christian Caron, University of North Carolina at Chapel Hill
Hailey Sherman, University of North Carolina at Chapel Hill

Draft for comments before submission, March 6, 2020

Appendix

Alternative ZINB models

In order to test the robustness of the effect of previous cumulative death sentences on predicting death sentences, we conduct a series of models with alternate specifications of variables and county filters. Across the tests, the effect of previous cumulative death sentences remains stable and statistically significant.

Given the variance in county population size, and thus total number of homicides and death sentences, we first run the full ZINB model on subsets of the data, based on both county population size and homicide counts. Table 5 lists the results specifically for the effect of previous cumulative death sentences by each county subset, and the full model results are presented in Tables 7 and 8. The stricter filters on counties do slightly decrease the value of the coefficient, but it remains generally consistent.

We also confirm the robustness of our analysis with additional variables and alternate measurements of variables. Table 6 contains the condensed results, with the full models in Tables 9-11. Based on the research of Jacobs, Carmichael, and Kent (2005) and Jacobs and

Carmichael (2002), we run models that include an interaction between racial threat and lynchings, the violent and property crime rates, a measure of evangelicalism, and an indicator about the partisanship of the legislature. Additionally, we substitute in other measures of variables, such as unemployment rather than poverty and different estimates of the lynching variable, to ensure that the effect is not due to the specific variables used in the full model. These results also provide evidence of the robustness of the effect of previous cumulative death sentences as the coefficients remain stable and statistically significant.

Table 5. Coefficient of Previous Cumulative Death Sentence in Limited Counties

	IRR for Cumulative Death Sentences _{t-1}	Robust Standard Error for Cumulative Death Sentences _{t-1}
Counties with Population Size > 10,000	1.08***	(-0.02)
Counties with Population Size > 25,000	1.07**	(-0.02)
Counties with Population Size > 50,000	1.07**	(-0.02)
Counties with Population Size > 100,000	1.07**	(-0.02)
Counties with Homicide Count > 4	1.07**	(-0.02)
Counties with Homicide Count > 9	1.07***	(-0.02)
Counties with Homicide Count > 24	1.06***	(-0.02)
Counties with Homicide Count > 49	1.05**	(-0.02)

***p < 0.001, **p < 0.01, *p < 0.05

Table 6. Coefficient of Previous Cumulative Death Sentence in Alternate Models

County	IRR for Cumulative Death Sentences _{t-1}	Robust Standard Error for Cumulative Death Sentences _{t-1}
Full Model with Threat*Lynching Interaction	1.08***	(-0.02)
Full Model with Violent Crime Rate	1.08***	(-0.02)
Full Model with Violent Crime Rate and Property Crime Rate	1.08***	(-0.02)
Full Model with Percent Evangelical	1.08***	(-0.02)
Full Model with Republican Legislature	1.08***	(-0.02)
Model with Homicide Rate	1.08***	(-0.02)
Model with White Population	1.08**	(-0.03)
Model with Black Population	1.08**	(-0.02)
Model with Unemployment	1.08***	(-0.02)
Model with Culture	1.08***	(-0.02)
Model with Black Lynching (Alternate Measure)	1.08***	(-0.02)
Model with Black Lynching (<i>This Cruel War</i>)	1.08***	(-0.02)
Model with Total Lynching (<i>This Cruel War</i>)	1.08***	(-0.02)
Model with Percent Slave	1.08**	(-0.02)

***p < 0.001, **p < 0.01, *p < 0.05

Table 7. Model Results – Varying Population Sizes

	Counties with Population >10,000	Counties with Population >25,000	Counties with Population >50,000	Counties with Population >100,000
<i>Death Sentence Absence</i>				
Cumulative Death Sentences _{t-1} /10	0.03*** (-0.02)	0.04*** (-0.02)	0.04*** (-0.03)	0.06*** (-0.04)
Ln Population	0.82* (-0.06)	0.90 (-0.09)	0.97 (-0.12)	1.24 (-0.20)
Homicides _{t-1} /100	1.03 (-0.07)	1.01 (-0.07)	1.02 (-0.07)	0.99 (-0.07)
Racial Threat/100	0.84 (-0.67)	0.69 (-0.66)	0.76 (-0.86)	0.63 (-0.93)
Percent in Poverty/10	1.69** (-0.30)	1.74** (-0.35)	1.78* (-0.43)	1.73* (-0.46)
Citizen Ideology/100	4.26 (-4.62)	6.38 (-8.08)	3.85 (-5.52)	2.96 (-4.58)
Lynchings/100	1.29 (-2.43)	3.62 (-7.12)	4.49 (-9.88)	2.62 (-8.26)
Partisan Supreme Court Elections	0.71* (-0.10)	0.67* (-0.11)	0.60** (-0.11)	0.50** (-0.13)
Republican Governor	0.92 (-0.11)	0.91 (-0.12)	0.87 (-0.13)	0.83 (-0.16)
South	0.62** (-0.11)	0.72 (-0.15)	0.67 (-0.16)	0.76 (-0.21)
Constant	-0.75 (2.16)	-1.36 (2.57)	6.09 (4.62)	3.63 (1.25)
Year FE	Yes	Yes	Yes	Yes
<i>One or more Death Sentences</i>				
Cumulative Death Sentences _{t-1} /10	1.08*** (-0.02)	1.07** (-0.02)	1.07** (-0.02)	1.07** (-0.02)
Ln Population	2.00*** (-0.10)	2.03*** (-0.11)	2.01*** (-0.12)	2.07*** (-0.13)
Homicides _{t-1} /100	0.99 (-0.02)	0.99 (-0.02)	1.00 (-0.02)	0.99 (-0.02)
Racial Threat/100	2.77* (-1.34)	2.78 (-1.46)	2.98* (-1.64)	2.86 (-1.68)
Percent in Poverty/10	1.33* (-0.17)	1.34* (-0.18)	1.35* (-0.20)	1.40* (-0.21)
Citizen Ideology/100	0.33 (-0.22)	0.33 (-0.23)	0.29 (-0.22)	0.23* (-0.17)

Lynchings/100	7.40** (-5.59)	7.99** (-6.03)	7.80** (-5.98)	8.26** (-6.49)
Partisan Supreme Court Elections	0.98 (-0.08)	0.97 (-0.08)	0.96 (-0.08)	0.94 (-0.09)
Republican Governor	1.00 (-0.06)	1.00 (-0.06)	0.98 (-0.06)	0.96 (-0.07)
South	1.00 (-0.11)	1.01 (-0.12)	1.00 (-0.12)	0.97 (-0.12)
Constant	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
Year FE	Yes	Yes	Yes	Yes
Total obs.	19,721	9,603	4,537	2,532
Nonzero obs.	3,552	2,683	1,803	1,217
Log pseudolikelihood	-10596.72	-7424.925	-4714.66	-3150.78

***p < 0.001, **p < 0.01, *p < 0.05 Notes: Reported Values are Incidence Rate Ratios. Robust standard errors clustered by county in parentheses. Each variable is rescaled by the factor indicated in order to generate coefficients that can be more easily interpreted. Models specified with year fixed effects.

Table 8. Model Results – Varying Homicide Counts

	Counties with Homicide Count > 4	Counties with Homicide Count > 9	Counties with Homicide Count > 24	Counties with Homicide Count > 49
<i>Death Sentence Absence</i>				
Cumulative Death Sentences _{t-1} /10	0.06*** (-0.03)	0.06*** (-0.03)	0.16*** (-0.07)	0.21*** (-0.08)
Ln Population	0.99 (-0.07)	0.99 (-0.07)	1.01 (-0.07)	1.12 (-0.10)
Homicides _{t-1} /100	1.13 (-0.17)	1.24 (-0.22)	1.36 (-0.41)	0.77 (-0.34)
Racial Threat/100	2.86 (-3.46)	0.70 (-1.06)	0.04 (-0.10)	0.67 (-1.45)
Percent in Poverty/10	2.03*** (-0.42)	2.25*** (-0.51)	2.05* (-0.70)	1.04 (-0.57)
Citizen Ideology/100	4.92 (-7.93)	5.00 (-8.07)	2.12 (-3.74)	0.93 (-2.70)
Lynchings/100	3.89 (-8.63)	4.60 (-11.86)	0.02 (-0.07)	0.00 (0.00)
Partisan Supreme Court Elections	0.65* (-0.14)	0.53* (-0.14)	0.55 (-0.30)	1.09 (-0.68)
Republican Governor	0.83 (-0.16)	0.68 (-0.16)	0.45** (-0.14)	0.38* (-0.17)
South	0.74 (-0.20)	0.82 (-0.26)	0.85 (-0.46)	0.83 (-0.70)
Constant	10.61 (8.36)			
Year FE	Yes	Yes	Yes	Yes
<i>One or more Death Sentences</i>				
Cumulative Death Sentences _{t-1} /10	1.07** (-0.02)	1.07*** (-0.02)	1.06*** (-0.02)	1.05** (-0.02)
Ln Population	2.07*** (-0.13)	2.07*** (-0.13)	2.09*** (-0.18)	2.02*** (-0.28)
Homicides _{t-1} /100	0.99 (-0.02)	0.99 (-0.02)	0.98 (-0.02)	0.99 (-0.02)
Racial Threat/100	2.90 (-1.62)	2.30 (-1.38)	2.17 (-1.56)	2.23 (-1.68)
Percent in Poverty/10	1.38* (-0.18)	1.39* (-0.19)	1.34 (-0.22)	1.35 (-0.27)
Citizen Ideology/100	0.27 (-0.20)	0.211* (-0.15)	0.196* (-0.14)	0.12** (-0.09)

Lynchings/100	8.06** (-6.03)	7.95** (-6.07)	9.05** (-7.01)	3.15 (-3.43)
Partisan Supreme Court Elections	0.96 (-0.09)	0.96 (-0.09)	0.98 (-0.13)	1.03 (-0.16)
Republican Governor	0.96 (-0.07)	0.93 (-0.07)	0.88 (-0.07)	0.89 (-0.09)
South	0.98 (-0.12)	0.91 (-0.12)	0.88 (-0.12)	0.78 (-0.11)
Constant				

Year FE	Yes	Yes	Yes	Yes
Total obs.	78,650	49,205	28,816	16,371
Nonzero obs.	4,968	4,561	3,967	3,261
Log pseudolikelihood	-17082.94	-14958.96	-12291.47	-9535.47

***p < 0.001, **p < 0.01, *p < 0.05 Notes: Reported Values are Incidence Rate Ratios. Robust standard errors clustered by county in parentheses. Each variable is rescaled by the factor indicated in order to generate coefficients that can be more easily interpreted. Models specified with year fixed effects.

Table 9. Model Results – Additional Variables

	Full Model with Threat* Lynching Interaction	Full Model with Violent Crime Rate	Full Model with Violent Crime Rate and Property Crime Rate	Full Model with Percent Evangelical	Full Model with Republican Legislature
<i>Death Sentence</i>					
<i>Absence</i>					
Cumulative Death Sentences _{t-1} /10	0.03*** -0.02	0.03*** -0.02	0.04*** -0.02	0.03*** -0.01	0.03*** -0.02
Ln Population	0.79** -0.06	0.76*** -0.06	0.75*** -0.06	0.80** -0.06	0.79*** -0.06
Homicides _{t-1} /100	1.05 -0.07	1.05 -0.07	1.05 -0.07	1.05 -0.07	1.04 -0.07
Racial Threat/100	0.34 -0.35	1.01 -0.78	1.12 -0.86	0.40 -0.33	0.80 -0.62
Percent in Poverty/10	1.72** -0.32	1.68** -0.29	1.68** -0.30	1.70*** -0.27	1.72** -0.30
Citizen Ideology/100	4.29 -4.56	6.41 -6.65	6.81 -6.77	6.02 -6.41	3.51 -3.49
Lynchings/100	0.00 0.00	1.75 -3.40	1.70 -3.49	1.52 -2.80	2.20 -4.32
Partisan Supreme Court Elections	0.70* -0.10	0.79 -0.13	0.77 -0.13	0.77 -0.11	0.74* -0.11
Republican Governor	0.95 -0.11	0.96 -0.11	0.95 -0.11	0.96 -0.12	0.91 -0.10
South	0.62** -0.11	0.59** -0.11	0.58** -0.11	0.65* -0.12	0.65* -0.12
Racial Threat * Lynching	2.45E+12 -4.00E+13				
Violent Crime Rate/100		0.96 -0.04	0.98 -0.05		
Property Crime Rate			1.00 0.00		
Percent Evangelical				1.00 -0.01	

Republican Legislature					2.77 -1.80
Constant	4.86*** (1.26)	4.35*** (1.19)	4.28*** (1.18)	4.65*** (1.14)	3.57** 1.25
Year FE	Yes	Yes	Yes	Yes	Yes
<hr/>					
<i>One or more Death Sentences</i>					
<hr/>					
Cumulative Death Sentences _{t-1} /10	1.08*** (-0.02)	1.08*** (-0.02)	1.08*** (-0.02)	1.08*** (-0.02)	1.08*** (-0.02)
Ln Population	0.99 (-0.02)	0.99 (-0.02)	0.99 (-0.02)	0.99 (-0.02)	1.00 (-0.02)
Homicides _{t-1} /100	1.98*** (-0.10)	1.96*** (-0.09)	1.95*** (-0.09)	2.12*** (-0.10)	1.98*** (-0.09)
Racial Threat/100	2.53 (-1.63)	2.74* (-1.29)	2.72* (-1.27)	2.02 (-0.98)	2.69* (-1.27)
Percent in Poverty/10	1.33* (-0.18)	1.34* (-0.17)	1.34* (-0.18)	1.32* (-0.15)	1.32* (-0.17)
Citizen Ideology/100	0.33 (-0.22)	0.28* (-0.18)	0.33 (-0.20)	0.41 (-0.25)	0.27* (-0.17)
Lynchings/100	1.06 (-7.76)	6.63* (-5.20)	7.34* (-6.15)	4.31* (-3.10)	8.70** (-6.72)
Partisan Supreme Court Elections	0.98 (-0.08)	1.06 (-0.10)	1.07 (-0.10)	0.97 (-0.08)	0.99 (-0.08)
Republican Governor	1.01 (-0.06)	1.00 (-0.06)	1.00 (-0.06)	1.02 (-0.06)	1.00 (-0.06)
South	0.99 (-0.11)	0.92 (-0.11)	0.90 (-0.11)	0.95 (-0.10)	0.97 (-0.11)
Racial Threat * Lynching	11.15 (-92.18)				
Violent Crime Rate/100		1.04 (-0.02)	1.03 (-0.03)		
Property Crime Rate			1.00 (0.00)		
Percent Evangelical				1.01* (0.00)	

Republican Legislature					1.38 (-0.56)
Constant	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
Year FE	Yes	Yes	Yes	Yes	Yes
Total obs.	102,070	102,070	102,076	84,148	98,257
Nonzero obs.	5,057	5,057	5,059	4,594	5,028
Log pseudolikelihood	-17635.48	-17616.23	-17626.19	-15873.11	-17502.12

***p < 0.001, **p < 0.01, *p < 0.05 Notes: Reported Values are Incidence Rate Ratios. Robust standard errors clustered by county in parentheses. Each variable is rescaled by the factor indicated in order to generate coefficients that can be more easily interpreted. Models specified with year fixed effects.

Table 10. Model Results – Alternative Variables

	Model with Homicide Rate	Model with White Population	Model with Black Population	Model with Unemploy- ment	Model with Culture
<i>Death Sentence Absence</i>					
Cumulative Death Sentences _{t-1} /10	0.03*** (-0.02)	0.03*** (-0.02)	0.03*** (-0.02)	0.03*** (-0.02)	0.03*** (-0.02)
Ln Population	0.76*** (-0.06)	0.81*** (-0.05)	0.79*** (-0.05)	0.73*** (-0.05)	0.76*** (-0.05)
Homicides _{t-1} /100		1.06 (-0.07)	1.05 (-0.07)	1.08 (-0.07)	1.06 (-0.07)
Racial Threat/100	1.59 (-1.25)			0.85 (-0.68)	0.66 (-0.52)
Percent in Poverty/10	1.77** (-0.33)	1.98*** (-0.38)	1.80*** (-0.32)		1.73** (-0.30)
Citizen Ideology/100	2.53 (-2.64)	4.12 (-4.34)	4.92 (-5.28)	3.27 (-3.30)	4.40 (-4.65)
Lynchings/100	2.64 (-5.57)	1.21 (-2.37)	4.21 (-8.50)	3.03 (-5.61)	1.98 (-3.63)
Partisan Supreme Court Elections	0.69** (-0.10)	0.72* (-0.10)	0.71* (-0.10)	0.76* (-0.11)	0.68** (-0.09)
Republican Governor	0.98 (-0.11)	0.98 (-0.11)	0.97 (-0.11)	0.93 (-0.11)	0.98 (-0.11)
South	0.64* (-0.11)	0.62** (-0.11)	0.67* (-0.13)	0.65* (-0.11)	
Homicide Rate	0.00*** (-0.01)				
White Population		1.01* (-0.01)			
Percent Black			0.99 (-0.01)		
Percent Black * Percent Black			1.00 (0.00)		
Unemployment				1.03 (-0.04)	
Culture					0.87** (-0.04)
Constant	4.32*** (1.23)	2.53 (1.30)	3.80** (1.16)	4.81*** (1.23)	5.13*** (1.16)

Year FE	Yes	Yes	Yes	Yes	Yes
<i>One or more</i>					
<i>Death Sentences</i>					
Cumulative Death Sentences _{t-1} /10	1.08*** (-0.02)	1.08** (-0.03)	1.08** (-0.02)	1.08*** (-0.02)	1.08*** (-0.02)
Ln Population	1.97*** (-0.10)	2.06*** (-0.09)	2.02*** (-0.09)	1.96*** (-0.09)	2.03*** (-0.09)
Homicides _{t-1} /100		0.99 (-0.02)	0.99 (-0.02)	1.00 (-0.02)	0.99 (-0.02)
Racial Threat/100	2.99* (-1.44)			3.03* (-1.43)	2.55* (-1.21)
Percent in Poverty/10	1.25 (-0.16)	1.43** (-0.19)	1.39** (-0.17)		1.33* (-0.16)
Citizen Ideology/100	0.26* (-0.17)	0.30 (-0.20)	0.30 (-0.20)	0.26* (-0.17)	0.48 (-0.30)
Lynchings/100	9.06** (-7.15)	7.00* (-5.43)	8.94** (-7.20)	7.65** (-5.62)	5.79* (-4.27)
Partisan Supreme Court Elections	0.99 (-0.08)	0.99 (-0.08)	0.99 (-0.08)	1.00 (-0.08)	0.97 (-0.07)
Republican Governor	1.03 (-0.06)	1.03 (-0.06)	1.03 (-0.06)	1.00 (-0.06)	1.04 (-0.06)
South	0.98 (-0.11)	1.01 (-0.12)	0.96 (-0.12)	1.03 (-0.12)	
Homicide Rate	0.80 (-0.45)				
White Population		1.00 (0.00)			
Percent Black			1.02* (-0.01)		
Percent Black * Percent Black			1.00** (0.00)		
Unemployment				1.05* (-0.02)	
Culture					1.04 (-0.03)
Constant	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
Year FE	Yes	Yes	Yes	Yes	Yes
Total obs.	102,064	102,070	102,003	102,070	97,352
Nonzero obs.	5,057	5,057	5,034	5,057	4,995

Log	-17600.41	-17652.45	-17577.23	-17639.47	-17347.59
pseudolikelihood					

***p < 0.001, **p < 0.01, *p < 0.05 Notes: Reported Values are Incidence Rate Ratios. Robust standard errors clustered by county in parentheses. Each variable is rescaled by the factor indicated in order to generate coefficients that can be more easily interpreted. Models specified with year fixed effects.

Table 11. Model Results – Alternative Lynching Variables

	Model with Black Lynching (Alternate Measure)	Model with Black Lynching (<i>This Cruel War</i>)	Model with Total Lynching (<i>This Cruel War</i>)	Model with Percent Slave
<i>Death Sentence Absence</i>				
Cumulative Death Sentences _{<i>t-1</i>}	0.03*** (-0.02)	0.03*** (-0.02)	0.03*** (-0.02)	0.03*** (-0.02)
Ln Population	0.78*** (-0.05)	0.78*** (-0.06)	0.77*** (-0.05)	0.78*** (-0.06)
Homicides _{<i>t-1</i>}	1.05 (-0.07)	1.05 (-0.07)	1.05 (-0.07)	1.05 (-0.07)
Racial Threat	0.70 (-0.52)	0.72 (-0.54)	0.79 (-0.61)	0.60 (-0.45)
Percent in Poverty	1.66** (-0.26)	1.67** (-0.27)	1.64** (-0.27)	1.71*** (-0.27)
Citizen Ideology	4.16 (-4.39)	4.32 (-4.53)	4.68 (-4.92)	3.75 (-3.92)
Lynchings	1.01 (-0.02)			
Partisan Supreme Court Elections	0.73* (-0.10)	0.73* (-0.10)	0.73* (-0.10)	0.72* (-0.10)
Republican Governor	0.96 (-0.11)	0.95 (-0.11)	0.95 (-0.11)	0.96 (-0.11)
South	0.61** (-0.11)	0.62** (-0.11)	0.61** (-0.11)	0.68 (-0.14)
Black Lynchings (<i>This Cruel War</i>)		1.00 (-0.02)		
Total Lynchings (<i>This Cruel War</i>)			1.01 (-0.01)	
Percent Slavery				0.89 (-0.39)
Constant	4.40*** (1.20)	4.35*** (1.20)	4.34*** (1.20)	4.52*** (1.21)
Year FE	Yes	Yes	Yes	Yes
<i>One or more Death Sentences</i>				
Cumulative Death Sentences _{<i>t-1</i>}	1.08*** (-0.02)	1.08*** (-0.02)	1.08*** (-0.02)	1.08** (-0.03)

Ln Population	1.99*** (-0.09)	1.98*** (-0.09)	1.97*** (-0.09)	2.00*** (-0.10)
Homicides _{t-1}	1.00 (-0.02)	0.99 (-0.02)	1.00 (-0.02)	0.99 (-0.02)
Racial Threat	2.78* (-1.274)	2.79* (-1.306)	3.04* (-1.471)	2.44 (-1.127)
Percent in Poverty	1.34** (-0.15)	1.34** (-0.15)	1.32* (-0.16)	1.39** (-0.16)
Citizen Ideology	0.31 (-0.21)	0.33 (-0.22)	0.35 (-0.23)	0.29 (-0.19)
Lynchings	1.03*** (-0.01)			
Partisan Supreme Court Elections	0.99 (-0.08)	0.99 (-0.08)	0.99 (-0.08)	0.99 (-0.08)
Republican Governor	1.02 (-0.06)	1.01 (-0.06)	1.02 (-0.06)	1.02 (-0.06)
South	1.00 (-0.11)	0.99 (-0.11)	0.99 (-0.11)	1.04 (-0.13)
Black Lynchings (<i>This Cruel War</i>)		1.02** (-0.01)		
Total Lynchings (<i>This Cruel War</i>)			1.02** (-0.01)	
Percent Slavery				1.11 (-0.28)
Constant	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
Year FE	Yes	Yes	Yes	Yes
Total obs.	102,070	102,070	102,070	102,070
Nonzero obs.	5,057	5,057	5,057	5,057
Log pseudolikelihood	-17638.7	-17640.09	-17640.65	-17654.2

***p < 0.001, **p < 0.01, *p < 0.05 Notes: Reported Values are Incidence Rate Ratios. Robust standard errors clustered by county in parentheses. Each variable is rescaled by the factor indicated in order to generate coefficients that can be more easily interpreted. Models specified with year fixed effects.

Constructing the homicide database

The homicide database was constructed using two data sources—both the Centers for Disease Control (CDC) and the Federal Bureau of Investigation (FBI) provide counts of homicides over time for the entire nation. The CDC reports are based on death certificates filed with county health departments whereas the FBI numbers come from local police agencies. CDC numbers are consistently higher than FBI numbers because they list all victims of homicide, whether or not the police have solved the crime. Note that the CDC numbers are organized by the county of residence of the victim, whereas FBI numbers refer to the county of the crime. In spite of differences in reporting practices, the two series track closely with each other; populous counties have more homicides, no matter which indicator is used. A regression predicting the CDC number by the FBI number (with no constant term) shows that there are 1.0225 CDC homicides for every homicide recorded by FBI; the R^2 for this one-variable equation is 0.9785.

The two series do not cover exactly the same time periods and have some gaps, so we take advantage of the high correlation between the two indicators to generate a single combined indicator based on the CDC numbers where available. Our final database is constructed as follows, with each estimate being used only if an estimate or observed value from the previous step was not available:

1. Direct observations based on CDC reports by county by year; approximately 37 percent.
2. Linear interpolation for individual missing observations where CDC observations are available for that county in contiguous years; approximately 1 percent.
3. Regression estimates of the CDC number based on FBI data; approximately 51 percent, particularly for counties with populations under 100,000 in the period after 1989, when the CDC stopped reporting such numbers.

4. Estimates of county values based on the typical share of total state homicides reported by the CDC, particularly for the period of 1989 through 2004. During this time the CDC reported state totals but not for individual counties, except for the largest counties. This affects approximately 4 percent of all observations.
5. Estimates of county values based on the typical share of total US homicides reported by the CDC, for the period of 2005 to 2019; this affects approximately 7 percent of all observations.