How Much Do Minority Lives Matter?

Jeff Witmer
Oberlin College

Journal of Statistics Education Volume 23, Number 2 (2015),
www.amstat.org/publications/jse/v23n2/witmer.pdf

Key Words: Simpson's Paradox.

Abstract

There are many well-known data sets that can be used to illustrate Simpson's Paradox. The Stand Your Ground data presented here shows Simpson's Paradox. In these data, race plays the key role – and not in the way that some students expect.

1. Introduction

Late in the fall of 2014, the nation's conscience was shaken when a grand jury did not return an indictment in the Michael Brown case in Ferguson, Missouri, and just a few days later, the same thing happened in the Eric Garner case in New York City. As protests arose around the country, many of us considered how the teaching of statistics might inform the national dialogue on race in America. Although it is not possible to fully dissect either the Brown or the Garner case, we can look at patterns in data from the criminal justice system and look for evidence of racism. This kind of analysis has been done by others over the years, for example, see Radelet (1981) or Radelet and Pierce (2011). The pattern of racism found in such analyses is still present today, as this paper illustrates.

In the next section. I present data related to the Stand Your Ground law in Florida that have a simple structure. There is a yes/no response variable and there are two predictor variables: the race of the defendant and the race of the victim. People often expect that racism within the court system will be evident in minority defendants having high conviction rates in comparison to white defendants. It turns out that what matters much more than race of the defendant is the race of the victim. Controlling for victim's race versus ignoring victim's race gives an example of Simpson's Paradox.

2. Stand Your Ground

In 2012, Trayvon Martin was shot to death by George Zimmerman in Sanford, Florida. When Zimmerman stood trial, he invoked Florida's Stand Your Ground law in his defense. After he was acquitted, the Tampa Bay Times created a webpage on which they presented data on cases in Florida in which the Stand Your Ground law was used by the accused person as part of their defense strategy; see http://www.tampabay.com/stand-your-ground-law/data. I looked at the 220 cases (both fatal and nonfatal) from this database for which full information is available (e.g., the case is not listed as "pending") and recorded (1) the race of the person accused of assault, (2) the race of the victim, and (3) the outcome (convicted or not convicted). Note that the website gives three races: white, black, and Hispanic; I reduced this to white and minority. The data set can be found here: http://www.amstat.org/publications/jse/v23n2/Simpson_Stand_Ground_2015.csv
The description of the data set can be found here:
http://www.amstat.org/publications/jse/v23n2/Simpson_Stand_Ground_2015_description.txt

Table 1 shows the cross-classification of (1) and (3). Note that the conviction rate was higher for white defendants (0.344) than for minority defendants (0.326).

Table 1.

		Defendant's race	
		Minority	White
Convicted?	No	60	86
	Yes	29 (0.326)	45 (0.344)

The 220 total cases included 88 in which the victim was a minority. Among these 88 cases, the conviction rate of white defendants was well below that of minority defendants; see Table 2.

Table 2. Minority Victims

		Defendant's race	
		Minority	White
Convicted?	No	45	19
	Yes	19 (0.297)	5 (0.208)

There were 132 cases in which the victim was white. For this subset of cases the conviction rate of white defendants was slightly below that of minority defendants; see Table 3.

Table 3. White Victims

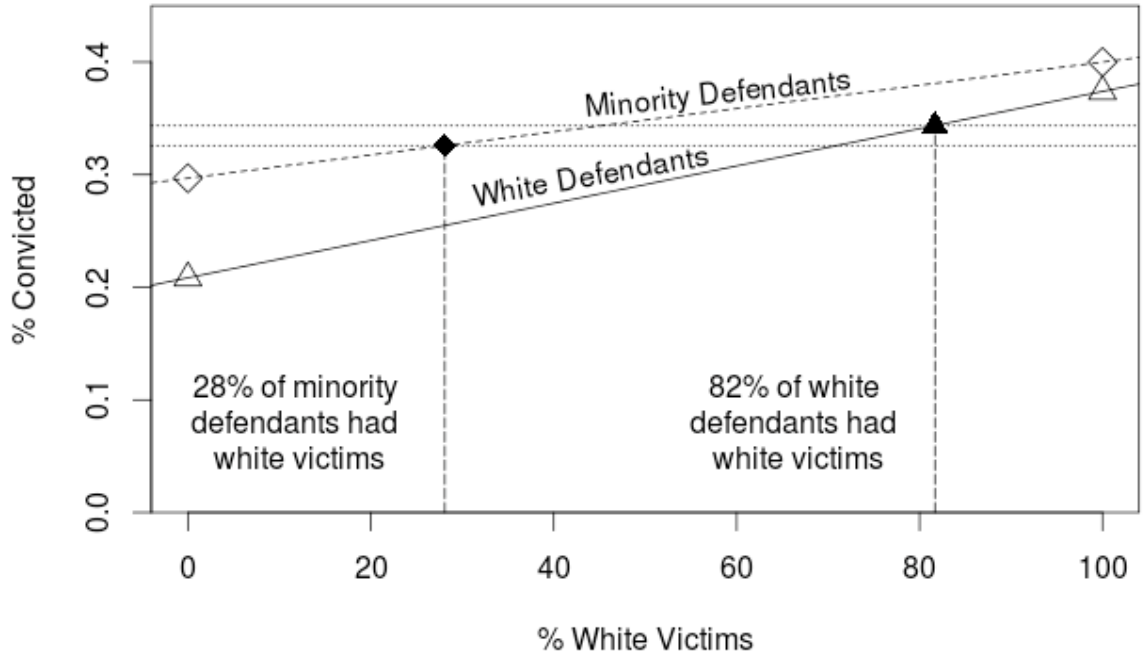
		Defendant's race	
		Minority	White
Convicted?	No	15	67
	Yes	10 (0.400)	40 (0.374)

This is an example of Simpson's Paradox: The white conviction rate is *lower* than the minority conviction rate when the victim is white and when the victim is not white, but the overall white conviction rate is *higher* than the minority conviction rate.

What matters is (2), the race of the victim. The combined conviction rate when the victim was white was 50/132 = 0.379 whereas the combined conviction rate when the victim was a minority was 24/88 = 0.273. Moreover, most of the white defendants had white victims, 107 of 131 (81.7%), while most of the minority defendants had minority victims, 64 of 89 (71.9%). Because whites defendants mostly had white victims and having a white victim greatly increases the chance of conviction, at first it appears that the conviction rate is higher among white defendants because of the defendant's race. But again, it is the race of the victim that is more important.

Figure 1 shows these relationships graphically. The diamonds represent minority defendants and the triangles represent white defendants. White victims are represented by the points at the right of the graph; minority defendants are represented by the points at the left of the graph. The open diamonds are above the open triangles, but the solid triangle – showing the overall conviction rate of white defendants – is higher than the solid diamond – showing the overall conviction rate of minority defendants.

Figure 1. A graphical display of Simpson's Paradox
White defendants have a lower conviction rate (open triangle versus open diamond) for either victim category, but a higher overall conviction rate (closed triangle versus closed diamond).



One can fit a logistic regression model in which conviction depends on the race of the defendant and see a positive coefficient: The fitted model is that the log odds of conviction equal -0.727 + 0.079*X, where X = 1 for white defendants and zero for minority defendants. A multiple logistic model shows a different picture: The fitted model is that the log odds of conviction equal -0.915 – 0.253*X1 + 0.624*X2, where X1 = 1 for white defendants and zero for minority defendants and X2 = 1 for white victims and 0 for minority victims.

The logistic fit shows that the race of the victim is a much stronger predictor of conviction than is the race of the defendant. Moreover, the coefficient on X1 is negative, meaning that being a white defendant decreases the odds of conviction, but the coefficient on X2 is positive, meaning that having a white victim increases the odds of conviction. This is evidence that crimes against whites are considered to be much more serious than crimes against minorities.

Helpful Hints for Teaching

This dataset illustrates Simpson's Paradox and underscores the prevalence of racism in the criminal justice system. When I present these data in my introductory courses, I start with Table 1 and ask whether this proves that white defendants are worse off than minority defendants. I then present Tables 2 and 3, placing them on either side of Table 1 on a whiteboard. I carefully show that "Table 2 + Table 3 = Table 1," lest students think that cases are added or deleted when creating these sub-tables. I then remark that if my team scores more points than your team in each half of a game, then my team wins, always. I ask the class how it can be that whites are worse off in each sub-table but better off overall. If needed, I nudge them toward the fact that whites tend to have white victims and the fact that having a white victim increases the chance of a conviction.

After establishing the paradox and presenting it graphically, with Figure 1, I ask the students what would happen if each cell count in Tables 2 and 3 were ten times larger. They quickly see that the percentages would all be the same and thus the paradox is not a function of sample size.

It is clear from the facial expressions of students, plus some comments, that this data set is disturbing to them, and thus I often end the class by apologizing for ruining their day. I also expect that they remember Simpson's Paradox because the example is so striking.

3. Conclusion

When someone is accused of a serious crime, the chance of conviction and the severity of the penalty depend on many things, including race. The race of the defendant matters, but what matters much more is the race of the victim. One wonders what would have happened in Ferguson and in New York if the victims had been white.

Acknowledgements

I thank the referees for helpful suggestions that have improved the presentation.

References

Radelet, M. (1981), "Racial Characteristics and the Imposition of the Death Penalty." *American Sociological Review,* 46, 918-27.

Radelet, M. and Pierce, G.L. (2011), "Race and Death Sentencing in North Carolina: 1980-2007." *North Carolina Law Review*. 89, 2119 – 60.

Schield, M. (2006), "Presenting Confounding Graphically Using Standardization." *STATS*, 46, 14 – 18.

Schneiter, K. and Symanzik, J. (2013), "An applet for the Investigation of Simpson's Paradox." *Journal of Statistics Education*, 21(1).

Jeff Witmer
Oberlin College
10 N. Professor St
Oberlin, OH 44074 USA
jeff.witmer@oberlin.edu
