

The Disproportion Statistic

N. Glenn Griesinger^{1,*}, Andrea J. Shelton², Jose Guardiola³ and Emiel W. Owens, Jr.⁴

¹ Department of Mathematics, Texas Southern University, Houston, Texas 77004, United States

² Department of Pharmacy Administration and Administrative Health Sciences, Texas Southern University, Houston, Texas 77004, United States

³ Department of Mathematics & Statistics, Texas A&M Corpus Christi, Corpus Christi, Texas 78412, United States

⁴ Department of Educational Administration and Foundations, Texas Southern University, Houston, Texas 77004, United States

Received: 23 Jun. 2018, Revised: 18 Sep. 2018, Accepted: 20 Sep. 2018

Published online: 1 May 2019

Abstract: We introduce a disparity measure, the disproportion statistic, suitable for use in the field of education. The reason for devising a disparity measure specifically for education stems from mounting evidence that extreme disciplinary actions result in greater numbers of repeat offenses and higher dropout rates. We derive the disproportion statistic then apply this statistic to disparities associated with zero tolerance policies as related to high school expulsions. We use the disproportion statistic to quantify the differences among ethnic groups and school settings. Patterns of disciplinary actions in four school settings, Non–Title I Non–Charter, Non–Title I Charter, Title I Non–Charter, and Title I Charter schools are examined. Analysis is based on the most recent data provided by the Office of Civil Rights. The importance of this research is that it is consistent with previous studies that show racial disparities due to zero tolerance policies. This research is unique because it additionally shows statistically significant differences in disparities among school settings as well.

Keywords: Disparity Measure, Zero Tolerance, School Discipline

1 Introduction

The term *disparities* indicates unfair or unjust differences. A statistical index of disparity is often used to compare the degree of racial/ethnic disparity among groups in such fields as healthcare [1], [9], [6] and education [2]. We present a new statistical measure of disparity not previously used in the field of education, the disproportion statistic. This rigorous measure is suitable for educational disparities because it deliberately does not weigh each subgroup by its fraction of the overall population, but averages the absolute differences between the proportion of members of each subgroup who receive a particular type of treatment and the proportion in the total population. To formulate the disproportion statistic, we apply a modification of the mean deviation of group rates to disparities in education [4]. We also theoretically derive its mean square error.

We focus on zero tolerance policy disparities among Blacks, Whites, Hispanics, Native Americans, Pacific Americans, and Asians. In schools, zero tolerance refers to the concept that certain types of disciplinary offenses will not be tolerated and will automatically result in suspension or expulsion. The school settings we focus on are Non–Title I Non–Charter schools, Non–Title I Charter schools, Title I Non–Charter schools, and Title I Charter schools. Non–Charter schools are traditional public schools. A Charter school is a publicly funded independent school. In order for a school to be classified as Title I, at least 40% of the students must be low income as define by the U. S. Department of Education.

The remainder of this paper is organized as follows. Section 2 defines the disproportion statistic and provides its derivation. Section 3 presents an application of the disproportion statistic. Section 4 evaluates the estimation quality of the disproportion statistic by theoretically deriving its mean square error. Conclusions are provided in Section 5.

* Corresponding author e-mail: Nancy.GlennGriesinger@tsu.edu

2 Disproportion Statistic Derivation

We implement a disparity measure to reflect the degree minorities receive disciplinary placements based on zero tolerance policies. This measure, the disproportion statistic, is defined as the fraction of the entire population that did not receive the most disciplinary placements relative to the group who did receive the most placements.

Assume the population is composed of k mutually exclusive subgroups. The data consists of the fraction [3],

$$p_i, \quad i = 1, \dots, k, \quad (1)$$

of individuals from each subgroup that received disciplinary placements. The fraction of the entire population the i^{th} group forms is γ_i :

$$\gamma_i = \frac{n_i}{N}, \quad (2)$$

where n_i is the number of individuals in subgroup i , and N is the population size. The data are placed in ascending order in terms of the size of p_i . The fraction of all members of the population receiving disciplinary action is:

$$\bar{p} = \sum_{i=1}^k \gamma_i p_i. \quad (3)$$

The disproportion statistic is bounded above by the value p_k , the fraction of individuals from the group who received the highest level of disciplinary action which is an expulsion. The disproportion statistic d_i for the i^{th} subgroup is the scaled difference between p_k and p_i , where $i = 1, \dots, (k-1)$ indexes the groups that did not receive the highest level of disciplinary action:

$$d_i = \gamma_i(p_k - p_i). \quad (4)$$

The disproportion statistic for the entire population is d :

$$d = \sum_{i=1}^{k-1} \gamma_i(p_k - p_i) = (1 - \gamma_k)p_k - \sum_{i=1}^{k-1} \gamma_i p_i = p_k - \bar{p}. \quad (5)$$

Equation (4) is used to plot d_i , $i \neq k$ for each ethnic group in each school setting. Equation (5) is used to compute the disproportion statistic for each school setting. Graphs and tables are presented in Section 3.

3 Disproportion Statistic Application

To determine if disparity is consistent over the four populations (school settings), we determine the d statistic for the four populations:

1. Non-Title I Non-Charter schools
2. Non-Title I Charter schools
3. Title I Non-Charter schools
4. Title I Charter schools

We also determine the statistic d_i , $i = 1 \dots 6$ for the six subgroups: Blacks, Whites, Hispanics, Native Americans, Pacific Americans, and Asians. Each statistic is scaled by a factor of 10^2 since values are close to zero.

Each table that follows represents one of the four school settings. The ethnic group that received the highest disciplinary placement served as a benchmark for each figure. The benchmark ethnic group is thus contained in the title of each figure. Specifically, tables contain the sample size of each of the 6 ethnic groups, the number of students in their particular ethnic group who were expelled due to zero tolerance, and each ethnic groups' corresponding disproportion statistic d_i , $i = 1 \dots 6$. Additionally, the overall d statistic is calculated for each table.

The scaled d statistic for Table 1 is $d = 0.0848$. Figure 1 depicts disparity of disciplinary placements based on zero tolerance policies received by Native Americans, who received the highest level of disciplinary placements, as compared to the five ethnic groups who received less than Native Americans in Non-Title I Non-Charter schools. Table 2 contains data for Non-Title I Charter schools. The scaled d statistic for Table 2 is $d = 0.2722$. Figure 2 demonstrates the disparity of disciplinary placements based on zero tolerance policies received by the same reference group, Native Americans, who received the highest level of disciplinary placements, as compared to the five other ethnic groups who received less than Native Americans in Non-Title I Charter schools. The scaled d statistic for Table 3 is $d = 0.0886$. Figure 3 shows disparity of disciplinary placements based on zero tolerance policies received by Native Americans, who received the highest level

Table 1: Table depicts **Non-Title I Non-Charter schools** total students enrolled, total students expelled due to zero tolerance policies, and the disproportion statistic d_i for each ethnic group. For the group that has the highest level of disciplinary placements, $d_i = 0$. The overall d statistic for the table is the sum of the d_i , $i = 1, \dots, (k - 1)$ for each ethnic group:

Ethnicity	Enrolled	Expelled	Disproportion, d_i
Black	1150984	1362	0.0078
White	5546608	4380	0.0631
Hispanic	1328719	1763	0.0068
Native Amer.	84937	150	0.0000
Pacific Amer.	36199	45	0.0002
Asian	431582	177	0.0068

Table 2: Table depicts **Non-Title I Charter schools** total students enrolled, total students expelled due to zero tolerance policies, and the disproportion statistic d_i for each ethnic group. For the group that has the highest level of disciplinary placements, $d_i = 0$. The overall d statistic for the table is the sum of the d_i , $i = 1, \dots, (k - 1)$ for each ethnic group:

Ethnicity	Enrolled	Expelled	Disproportion ($\times 10^2$)
Black	20299	22	0.0509
White	51515	54	0.1310
Hispanic	24530	12	0.0758
Native Amer.	1647	6	0.0000
Pacific Amer.	384	0	0.0012
Asian	3640	0	0.0130

Table 3: **Title I Non-Charter schools** total students enrolled, total students expelled due to zero tolerance policies, and the disproportion statistic d_i for each ethnic group. For the group that has the highest level of disciplinary placements, $d_i = 0$. The overall d statistic for the table is the sum of the d_i , $i = 1, \dots, (k - 1)$ for each ethnic group:

Ethnicity	Enrolled	Expelled	Disproportion ($\times 10^2$)
Black	678278	1477	0.0092
White	922809	1024	0.0447
Hispanic	1237654	2486	0.0236
Native Amer.	49808	129	0.0000
Pacific Amer.	19583	20	0.0010
Asian	146045	67	0.0102

Table 4: **Title I Charter schools** total students enrolled, total students expelled due to zero tolerance policies, and the disproportion statistic d_i for each ethnic group. For the group that has the highest level of disciplinary placements, $d_i = 0$. The overall d statistic for the table is the sum of the d_i , $i = 1, \dots, (k - 1)$ for each ethnic group:

Ethnicity	Enrolled	Expelled	Disproportion ($\times 10^2$)
Black	45526	101	0.0000
White	28632	28	0.0293
Hispanic	41492	58	0.0281
Native Amer.	1879	4	0.0001
Pacific Amer.	332	0	0.0006
Asian	3102	2	0.0040

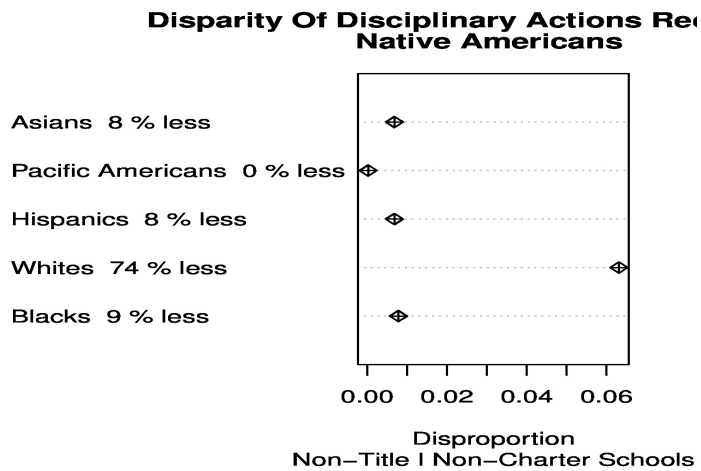


Fig. 1: Dot chart displays disparity of disciplinary placements based on zero tolerance policies received by Native Americans, who received the highest level of disciplinary placements, as compared to the five other ethnic groups who received less than Native Americans in *Non-Title I Non-Charter* schools.

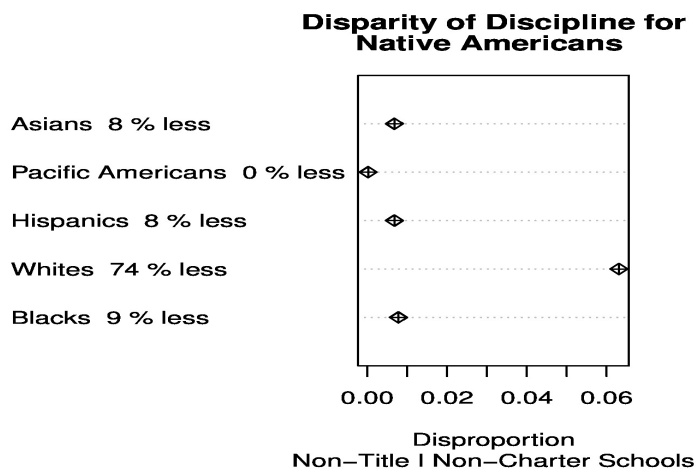


Fig. 2: Dot chart displays disparity of disciplinary placements based on zero tolerance policies received by Native Americans, who received the highest level of disciplinary placements, as compared to the five other ethnic groups who received less than Native Americans in *Non-Title I Charter* schools.

of disciplinary placements, as compared to the five other ethnic groups who received less than Native Americans in Title I Non-Charter schools.

Table 4 contains data for Title I Charter schools. The scaled d statistic for Table 4 is $d = 0.0623$. Figure 4 highlights disparity of disciplinary placements based on zero tolerance policies received by Blacks, who received the highest level of disciplinary placements, as compared to the five other ethnic groups who received less than Blacks in Title I Charter schools.

To test the disproportion statistics, d from the four tables, we perform a test of equality for multiple proportions. The null hypothesis is all proportions are equal, and the alternative is at least one proportion is different. The test statistic is d . The p -value for the given data set is 3.432×10^{-24} . Therefore, the null hypothesis is rejected at the $\alpha = .01$ level

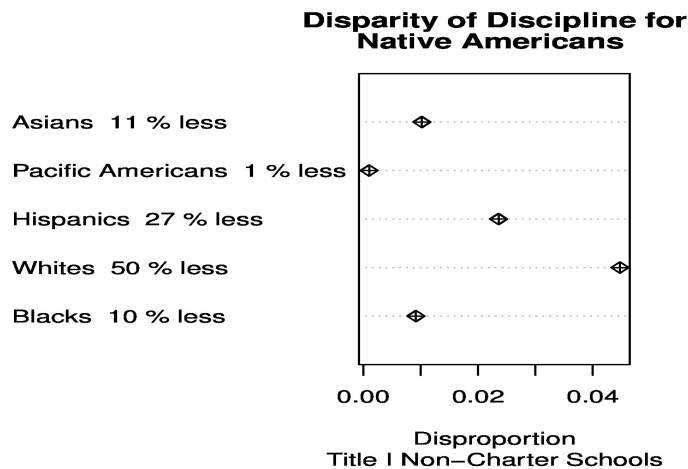


Fig. 3: Dot chart displays disparity of disciplinary placements based on zero tolerance policies received by Native Americans, who received the highest level of disciplinary placements, as compared to the five other ethnic groups who received less than Native Americans in *Title I Non-Charter* schools.

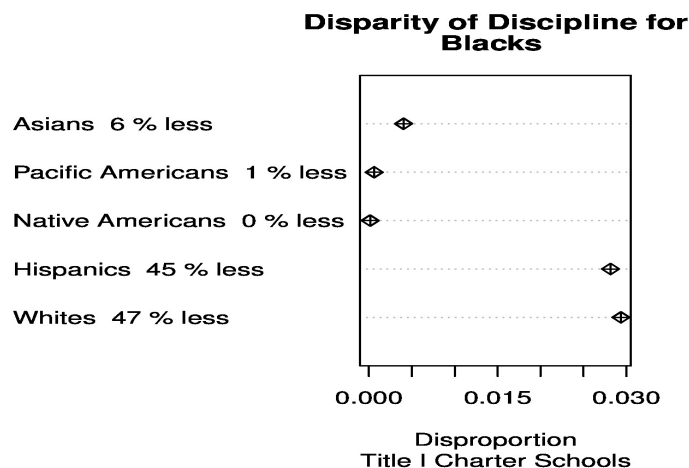


Fig. 4: Dot chart displays disparity of disciplinary placements based on zero tolerance policies received by Blacks, who received the highest level of disciplinary placements, as compared to the five other ethnic groups who received less than Blacks in *Title I Charter* schools.

of significance. Conclude that not all school systems are equal with respect to disparity. Therefore, it is important to determine which school systems are different.

We use the *R* function *pairwise.prop.test* to calculate pairwise comparisons between pairs of the four *d* statistics, then use the Holm procedure [5] to adjust the *p-values* for multiple testing. The function *pairwise.prop.test* enables us to simultaneously test pairs of proportions. From Table 5 conclude that all pairs of school settings have statistically significant differences with respect to disparity due to zero tolerance policies.

4 Mean Square Error of Disproportion Statistic

Since we can usually apply more than one disparity measure in a particular situation, a difficulty that may arise is the task of choosing the best estimator. Therefore, we need some criteria such as the mean square error to evaluate the quality of

Table 5: Pairwise Comparisons of disparity between schools, p-value, and statistical significance.

Comparison	p-value	significant
d_{table1}, d_{table2}	2.000×10^{-1}	yes
d_{table1}, d_{table3}	0.002×10^0	yes
d_{table1}, d_{table4}	2.000×10^{-16}	yes
d_{table2}, d_{table3}	2.000×10^{-16}	yes
d_{table2}, d_{table4}	2.000×10^{-16}	yes
d_{table3}, d_{table4}	2.000×10^{-16}	yes

an estimator. The mean square error of an estimator $\hat{\theta}$ of a parameter θ is defined as the expected value:

$$\text{MSE}(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]. \quad (6)$$

Theorem 1. Let d be the disproportion statistic. Since d is defined to be a proportion, the mean square error of the estimator $\text{MSE}(\hat{d}) = \left(\frac{pq}{n}\right) \left(\frac{N-n}{N-1}\right)$. Where p represents the fraction of the entire population that did not receive the most disciplinary placements due to zero tolerance policies relative to the group who did receive the most placements, and $q = 1 - p$. The variable N represents the population size (total students enrolled in a school setting). The variable n represents subgroup size.

Proof. Let $d \in \mathbb{R}$, and let $\hat{d}$ be an unbiased estimator of d . The population mean of d is μ . Without loss of generality, $E[(\hat{d} - \mu)] = 0$ and the sample proportion is an unbiased estimator of the population proportion. From the definition of mean square error,

$$\begin{aligned}
 \text{MSE}(\hat{d}) &= E[(\hat{d} - d)^2] \\
 &= E[(\hat{d} - \mu + \mu - d)^2] \\
 &= E[(\hat{d} - \mu)^2 + 2(\hat{d} - \mu)(\mu - d) + (\mu - d)^2] \\
 &= E[(\hat{d} - \mu)^2] + E[2(\hat{d} - \mu)(\mu - d)] + E[(\mu - d)^2] \\
 &= \text{Var}[\hat{d}] + 2(\mu - d)E[(\hat{d} - \mu)] + (\mu - d)^2 \\
 &= \text{Var}[\hat{d}] \\
 &= \left(\frac{pq}{n}\right) \left(\frac{N-n}{N-1}\right).
 \end{aligned}$$

5 Conclusions

Prior research shows that there are disparities among ethnic groups regarding the enforcement of zero tolerance policies [7], [8]. Creating a fair and positive school environment and devising creative approaches to discipline are strategies being considered to replace the rigid punishments associated with zero tolerance policies. The importance of this research is that for one of the first times it is shown that there are statistically significant differences in disparities due to zero tolerance policies among school settings as well as disparities among ethnic groups.

To quantify these differences, we compute the disproportion statistic for each ethnic group and for each of four school settings, Non-Title I Non-Charter, Non-Title I Charter, Title I Non-Charter, and Title I Charter schools. Findings show disciplinary disparities based on zero tolerance policies are largest among Native Americans and Blacks, who receive the greatest level of disciplinary placements, and that of Whites who receive the least. Findings also show that all pairs of school settings have statistically significant differences with respect to disparity due to zero tolerance policies.

References

- [1] L. N. Borrell and M. Talih, A symmetrized Theil index measure of health disparities: an example using dental caries in U.S. children and adolescents, *Statistics in Medicine*, **30**(3), 277 – 290 (2011).
 - [2] C. Na and D. C. Gottfredson Police Officers in Schools: Effects on School Crime and the Processing of Offending Behaviors. *Justice Quarterly*. **30**(4), 619–650 (2013).
 - [3] J. L. Gastwirth, A graphical summary of disparities in health care and related summary measures, *Journal of Statistical Planning and Inference*, **137**, 1059 – 1065 (2007).
 - [4] K. G. Keppel, J. N. Percy, and D. Wagener, Trends in racial and ethnic specific rates for the health status indicators, *Statistical Notes*, 23, Center for Disease Control, Atlanta (2002).
 - [5] M. H. Kutner, C. Nachtsheim, and J. Neter, *Applied Linear Statistical Models*, McGraw–Hill (2004).
 - [6] T. H. McCormick and T. Tian, Latent demographic profile estimation in hard-to-reach groups, *The Annals of Applied Statistics*, **6**(4), 1795 – 1813 (2012).
 - [7] R. J. Skiba, The failure of zero tolerance, *Reclaiming Children & Youth*, **22**(4), 27 – 33 (2014).
 - [8] R. J. Skiba, C. Chung, M. Trachok, T. L. Baker, A. Sheya, and R. L. Hughes, Parsing disciplinary disproportionately: Contributions of infraction, student and school characteristics to out-of-school suspension and expulsion, *American Education Research Journal*, **51**(4), 640 – 670 (2014).
 - [9] Talih, Makram, A reference-invariant health disparity index based on Rényi difference, *The Annals of Applied Statistics*, **7**(2), 1217–1243 (2013).
-



N. Glenn Griesinger earned her PhD degree in Statistics from Rice University in Houston, Texas. She is an Associate Professor of Statistics at Texas Southern University. Her research interests are in the areas of nonparametric statistics, disparity measures, and pulmonology measures. She is a statistical methodologist who has published research articles in reputed national and international statistical journals and in bioinformatics journals. She is an editor of the journal *Annals of Pulmonology*. She was also appointed to the American Statistical Association Committee on Minorities in Statistics.



Andrea J. Shelton earned her PhD degree in Epidemiology at the University of Pittsburgh, Graduate School of Public Health. She is a professor in the College of Pharmacy and Health Sciences at Texas Southern University in Houston, Texas. She has done global research on malaria, conducted national studies of various chronic diseases, and coordinated a funded project on mental health awareness and suicide prevention. She has also written articles about challenges in U.S. public schools, including safety issues, support of STEM instruction, and students' perceptions of an effective classroom learning environment.



Jose Guardiola earned his PhD degree in Statistics from Baylor University in Waco, Texas. He is an Associate Professor of Mathematics at Texas A&M University - Corpus Christi at the Department of Mathematics and Statistics. His research interests are in the areas of spatial statistics, distribution theory, circular statistics and entropy. He is also a statistical consultant for local hospitals. He has published research articles in national and international statistics journals. He is also a frequent reviewer of the *Numeracy* journal from the University of South Florida.



Emiel W. Owens, Jr. earned his EdD degree at the University of Houston with an emphasis in Research Methods and Mathematics Education. He is currently a professor in the College of Education at Texas Southern University and teaches courses in Educational Research and Statistics. His current research interest is in modeling longitudinal data.