

## **Racial Reporting Bias and Child Maltreatment**

### **Sheila D. Ards**

Vice President for Community Development  
Benedict College  
and  
Associate Professor  
Humphrey Institute of Public Affairs  
University of Minnesota

### **Samuel L. Myers, Jr.**

Roy Wilkins Professor of Human Relations and Social Justice  
Humphrey Institute of Public Affairs  
University of Minnesota

### **Chanjin Chung**

Associate Director  
Agricultural Commodity Programs  
Cornell University

### **Allan Malkis**

Research Associate  
Roy Wilkins Center for Human Relations and Social Justice  
Humphrey Institute of Public Affairs  
University of Minnesota

### **Brian Hagerty**

Research Assistant  
Roy Wilkins Center for Human Relations and Social Justice  
Humphrey Institute of Public Affairs  
University of Minnesota

For presentation at the International Society for  
The Prevention of Child Abuse and Neglect  
VIII European Conference on Child Abuse and Neglect  
24-27 August 2001  
Istanbul, Turkey

## Abstract

The two central national data sets on child abuse and neglect offer conflicting visions of the racial composition of the abused and neglected populations. The National Child Abuse and Neglect Data System (the primary source of national information on abuse and neglected children known to States) shows that African Americans are overrepresented among reported and substantiated abuse and neglect cases. This data set consistently reveals a disproportionate number of black children among those who are reported to Child Protective Services. The second data set is the National Incidence Study of Child Abuse and Neglect, collected in 1980, 1986, and 1993, shows no apparent overrepresentation of children of color. The estimated abuse rates for whites and nonwhites are not different statistically. In this paper, we briefly review some explanations of how these alternative explanations can both be true. We point out that there is little evidence of reporting bias. Still, the measurement of reporting effects may be influenced by another type of bias, called aggregation bias. Through several examples and an application using NIS data, we show that there may be an appearance of *racial* reporting bias due to the aggregation of poor and nonpoor children.

## **Introduction**

African American children make up a disproportionate share of children in Child Protective Systems (CPS) across the nation. A 1990 National Child Abuse and Neglect System report showed that African American children represent 25 percent of the reported and substantiated child abuse and neglect cases, while these children represent only 15 percent of the national child population. The seeming overrepresentation of African American children in the CPS may suggest that African American children are more likely to be abused and neglect. The National Incidence Studies (1980, 1986, and 1993), however, suggest that this is not the case. According to these studies, there are no statistical differences in child maltreatment by race. The purpose of this paper is to examine these two disparate findings and to offer an explanation as to why there are more African American children in the CPS while there are no racial differences in the occurrence of child maltreatment.

The two central national data sets on child abuse and neglect offer conflicting visions of the racial composition of the abused and neglected populations. The National Child Abuse and Neglect Data System (NCANDS) is “the primary source of national information on abused and neglected children known to States” (U.S. Department of Health and Human Services, 1999). NCANDS shows that African Americans are overrepresented among reported and substantiated abuse and neglect cases. This data set consistently reveals a disproportionate number of black children among those who are reported to Child Protective Services. For example, in 1995, with 44 states reporting race/ethnicity of victims, 55 percent of child abuse and neglect victims were white and 27 percent were black (U.S. Department of Health and Human Services, 1997, pp. 2-9). Yet among children under 18 in the general population, 66 percent are white and 15 percent are black (U.S. Bureau of Census, 1996, Table No. 22). This suggests that black children are overrepresented and whites are underrepresented in the official child abuse population as recorded in NCANDS data.

The second data set is the National Incidence Study of Child Abuse and Neglect (NIS) collected in 1980, 1986, and 1993 (referred to here as NIS-1, NIS-2, and NIS-3, respectively). The results of the NIS studies show no apparent overrepresentation of children of color. The estimated abuse rates for whites and nonwhites are not different statistically (Sedlak, Hantman, and Schultz, 1997, February). The central distinction between the two data sets is that the NIS data include individual-level information on both reported and unreported incidents of abuse and neglect. The main statistics are provided on the number and rate of children who are abused by eliminating duplicated cases and assigning family characteristics to each child who is suspected of abuse or neglect. Thus, the unit of measurement is the individual child. There is a uniform definition of maltreatment even though the cases are collected from various counties across different states. The NCANDS data set is restricted to reported cases. Some information is only available for reported and substantiated cases. In addition, the NCANDS data sets by necessity rely on state delineations of abuse and neglect and state procedures for handling duplicated cases.

At first blush, then, one can speculate that the difference between the NIS and NCANDS measures of representation of African American children might be due to reporting (or substantiation) biases. But there may be other reasons why these data sets tell different stories about racial disparities.

In this paper, we briefly review some of these alternative explanations. We point out, with reference to our previous research, that there is little evidence of reporting bias. Still, the measurement of reporting effects may be influenced by another type of bias, called aggregation bias. This is the focus of our paper. Through several examples and an application using NIS data, we show that there may be an appearance of *racial* reporting bias due to the aggregation of poor and nonpoor children.

## Bias in Official Reports

An emerging literature points to potential biases in official reports. Brett Drake and Susan Zuravin (1998) note that many biases enter into the reporting of child maltreatment, which can lead to overestimates of child maltreatment among the poor. They identify four types of biases: visibility or “exposure bias,” labeling bias, reporting bias, and substantiation bias. They suggest that not enough evidence exists to determine whether the apparent disproportionate representation of poor children is due to these biases.

The first three authors of this paper previously have tested the reporting bias hypothesis. We used sample selection bias estimators to determine whether racial bias in reporting exists in the 1980 NIS data (Ards, et al., 1998). That work focused on whether the seeming overrepresentation of blacks in the NCANDS report data might be due to racial differences in reporting. A general definition of (racial) reporting bias is that identically situated abused or neglected children of different races are reported at different rates to child protective service workers by mandated professionals. A *report* is a case that becomes known to CPS that meets its definition of child maltreatment. Some cases are reported to a source such as a law enforcement agency, but the case is not officially reported to the CPS. Some cases are directly reported to the CPS by therapists, private physicians, or friends and family members. The *report rate* is the ratio of the number of reports to the number of *known* child maltreatment cases. The *probability of report* relates to the event that a *known* case becomes *reported* to the CPS. This definition of report rates and report probabilities makes it technically possible for measures to be inflated should the data collection design—like that of the NIS—fail to incorporate counts of *known* cases from people outside sampled agencies. For these reporters, the report rates (and probabilities of report) will be equal to one. Excluding these reporters from the analysis will understate the extent of child abuse and neglect, but including them will overstate the degree of reporting. This is a problem of *sample selection bias*. Using the NIS data

to measure report rates, and correcting for selection bias, we do not find consistent evidence of racial reporting bias. Our initial research, however, only looks at reporting bias, not other possible biases.

Robert Morton (1999), Child Welfare Institute president, points out that work such as ours looks only at one segment of the pathway from the occurrence of child maltreatment to final substantiation by CPS. He believes there may be substantiation bias. One can conceptualize substantiation bias along the same lines as reporting bias: *substantiation bias* occurs when reported allegations of maltreatment faced by children of different races are not substantiated at the same rate.

Drake and Zuravin identify quite starkly that one of the most important initial steps to recognition of abuse and neglect is taken by professionals. They posit that the poor—and by extension African Americans—may interact with a greater number of mandated professionals, and so are more likely to be known to these professionals. The *exposure bias* then arises from the fact that blacks (or the poor) are more likely to be found among those who are most likely to be reported. This conclusion parallels the research of Blumstein (1982) and others (Langan, 1985; Hindelang, 1978) who contend that in the criminal justice system, the apparent disparity in incarceration can be explained by the higher involvement of blacks in activities most likely to result in incarceration. Overrepresentation is much smaller when the comparison is made among the arrested population, researchers contend. Thus, exposure bias is another way of explaining racial disparities that point the policy lenses away from decision-makers or institutions—such as mandated professionals or the courts and prisons—toward victims and their families.

Another important bias well known in the sociological literature is *labeling*. A question that has been explored in the literature is whether the same case is viewed differently by different professionals. In other words, labeling bias results from the "same behavior . . . defined in any number of ways" (Macionis, 1987, Chapter 8, p. 215). One person may describe a case as involving physical discipline while another may say it involves physical abuse. Giovannoni and Becerra (1979) present one of the most thorough analyses of this phenomenon. They explored attitudinal differences

in classifying maltreatment by presenting case vignettes to professionals and lay people in a large metropolitan area. They were interested in seeing whether ethnic or professional differences play a role in identifying the severity of a situation on the welfare of a child. Their work showed that "community members saw most kinds of mistreatment as more serious than did professionals, and among professionals, lawyers especially dissented from the other groups, generally regarding mistreatment as less serious than the others did" (p. 240). Turbett and O'Toole (1980) showed that physicians were more likely to identify and report black children as maltreated than white children. Thus the background characteristics of the child played a role in whether the professional "saw" maltreatment. Our operationalization of the concept of labeling requires us to ask whether cases that were previously determined to be abuse or neglect signal a higher probability of report or substantiation for blacks than for whites. That is, is the same behavior assessed differently simply because a child is black and has previously been determined to be a victim of abuse or neglect? Examples that come to mind might be the presence of welts or strap marks. Case workers may believe that such bruises, when accompanied by prior abuse allegations, clearly "prove" that black children have been maltreated. Case workers have, in effect, *labeled* a particular form of discipline in the black community as maltreatment. If such behavior is not believed to be as prevalent in the white community, the presence of such marks and bruises would not immediately trigger the same response from case workers even when there is a prior allegation of abuse. We define labeling, then, to be the existence, given prior contact, of a lower threshold for blacks than for whites in determining maltreatment, indicating that blacks have in effect been singled out simply because they are black and prior victims. This interpretation of labeling is similar to that used in criminal justice theory where a prior conviction differentially labels certain activities as criminal.

Finally, there is the statistical problem of *aggregation bias*. This form of bias arises from combining substantively different units in a common aggregate. Examples include combining rural and urban cases; cases from differing regions; or cases that enter the system through differing

mechanisms. One example of aggregation bias that we explore in detail below arises from combining welfare and nonwelfare cases. Blacks are overrepresented among welfare recipients, who are more likely to be abused or neglected and thus more likely to be found among the abused or neglected. This bias can arise even if the rates of maltreatment reporting are the same among whites and blacks who receive welfare and those who do not. *Aggregation* can uncover racial disproportionality when none exists.

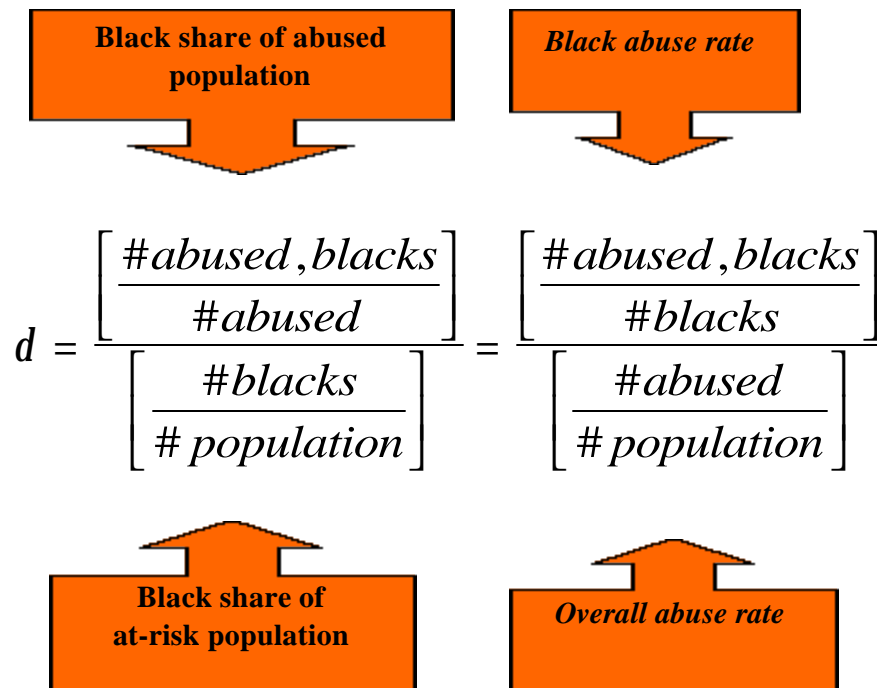
### Method

The thesis of this paper turns upon the notion of disproportionality. Disproportionality describes a situation in which the relative share of a single factor on two populations is substantially different.

For example, if we denote,  $d$  = disproportionality in child abuse, then we would define

disproportionality as  $d > 1$ , if blacks' share of the abused population exceeds their share of the at-risk population, where

$$d = \frac{\% \text{ blacks among abused}}{\% \text{ blacks among population}}$$



## The Problem of Aggregation Bias

Consider a world where there are 10,000 children (Table 1, Case 1). African Americans make up 1,000 of this hypothetical world; the other 9,000 are white. Suppose that the actual black abuse rate is 10%; the actual overall abuse rate is 10%. There is no disproportionality. If we find that the black (reported) abuse rate is 3% (= 30/1000), and the overall (reported) abuse rate is 1% (= 100/10,000), then the result is an observed disproportionality of 3:1. The measure of representation of African Americans apparently is biased upwards due to reporting bias. We should note, however, an important caveat: Case 1 does not take into account other intervening factors, such as characteristics of the child, characteristics of the alleged abuser, characteristics of the reporter, type of abuse, and factors related to the county or state where the abuse occurs, that might explain the difference. .

**Table 1: Case 1, Apparent Racial Bias in Reporting**

Population		Abused		Reported	
Black	White	Black	White	Black	White
1000	9000	100	900	30	70
Blacks = 10%		Blacks = 10%		Blacks = 30%	

**Table 2: Case 2, Aggregation Bias**

	Population		Abused		Reported	
	Black	White	Black	White	Black	White
Welfare	600	600	280	280	28	28
Non-Welfare	400	8400	20	420	2	42
Total	1000	9000	300	700	30	70

Now, consider a population in which blacks are 30 percent of the abused population and there is no disproportionality in abuse. In this example, blacks are overrepresented among a group, say the welfare population, that is more likely to be abused. From the figures given in Table 2, Case 2, we see that the abuse rate among blacks and whites who come into contact with the welfare office (46.6% = 280/600) is the same. Moreover, the report rates are the same for all subgroups (10%). Yet

in Case 2 precisely the same disproportionality is observed among reported cases as is found in Case 1. The key reason for the difference in the measured disproportionality is an underlying disparity in abuse between welfare- and nonwelfare-receiving persons. The nonwelfare abuse rate is only 5% ( $20/400 = 420/8400$ ), and is identical for blacks and whites. The aggregation of welfare and nonwelfare cases gives the impression that there is reporting bias when there is not. That is why it is necessary to measure reporting bias directly to be able to infer whether there may in fact be underlying racial differences in abuse.

But one rarely knows the actual level of abuse. Blacks may be overrepresented among the reported abuse cases even if there is no underlying difference between welfare and nonwelfare abuse rates. Instead, there may be a propensity to *report* welfare abuse cases more often than nonwelfare abuse cases, which could occur even if there were no racial differences in the report rates. The result could not be strictly termed racial reporting bias, but is rather another form of aggregation bias known as exposure bias. Case 3 (Table 3), makes this apparent.

**Table 3: Case 3, Exposure Bias**

	Population		Abused		Reported	
	Black	White	Black	White	Black	White
Welfare	600	600	60	60	28	28
Non-Welfare	400	8400	40	840	2	42
Total	1000	9000	100	900	30	70

In this instance, 10% of black children are abused; 10% of white children are abused. There is no disproportionality in abuse rates. There is, however, a racial disproportionality in reported abuse rates even though blacks and whites face the same report rates within groups. The disproportionality arises because welfare groups are more likely to be reported (46.6% vs 5%) and blacks are more likely to be on welfare. This is not racial reporting bias, but is a form of bias that might explain why blacks are disproportionately represented among the reported population.

## Data

In the previous examples, we showed that measurement issues may create the illusion of racial disproportionality in maltreatment when in fact there is none. We can also illustrate this point in a slightly different manner, using the National Incidence Study (NIS) data for 1980, 1986 and 1993 (NIS-1, NIS-2 and NIS-3).

The 1980 collection effort, mandated under Public Law 93-247 (1974), the 1986 collection effort, mandated under Public Law 98-457 (1984), and the 1993 collection effort, mandated under the Child Abuse Prevention, Adoption and Family Services Act of 1988 (Public Law 100-294, as amended) provide measures of both reported and unreported child maltreatment and thus can be used to estimate both maltreatment rates and reporting rates (Sedlak *et al.*, 1997, February, pp. 1-1, 1-2). A representative sample of counties was chosen. In each county, both child protective service agencies and non-child protective service agencies were selected for participation in the data collection. The non-CPS agencies included law enforcement agencies, medical services, and schools. Professionals posted in these agencies were asked to fill out forms indicating the presence of specific acts that meet standardized definitions of child abuse and neglect. The data collected in all three studies included the information from the sampled non-CPS agencies, cases directly reported to the CPS, and cases reported to the CPS from sources other than the sampled agencies, e.g., friends, neighbors, private physicians, and therapists in private practice (Sedlak *et al.*, 1997, February, pp. 1-3). In the NIS, mandated professionals report less than half of all cases known to the CPS. Most come from families, neighbors and friends.

Within the NIS, the unit of analysis is the child suspected of maltreatment. The NIS only records one incident: a child can only enter the data set once even if he or she is suspected of being abused or neglected again within a year. Moreover, if more than one agency reports the suspected case of maltreatment to the CPS—or if more than one person reports the suspected case of maltreatment to the mandated professional—duplicated cases are not counted in the final data set.

In these studies, child maltreatment was defined thus:

“...where, through purpose, acts, or marked inattention to the child's basic needs, behavior of a parent/substitute or other adult caretaker caused foreseeable and avoidable injury or impairment to a child or materially contributed to unreasonable prolongation or worsening of an existing injury or impairment.” (U.S. Department of Health and Human Services, 1981)

This definition overcomes several significant obstacles faced by other data collection efforts. It provides a uniform definition of child maltreatment across disciplines and across states. It allows child maltreatment cases to be compared across states and it delineates who can be counted as a perpetrator and what actions constitute maltreatment.

Although the aims of NIS-1, NIS-2, and NIS-3 are the same, variations occur in the methods of selecting the sentinels for the study, the time frames for data collection, and the expanded definition of child maltreatment. The effect of these differences is that the weights used for the analysis are sufficiently complicated to make simple replication of the 1980 analysis extremely time consuming. Separate programs must be written and performed for each year.

We compute maltreatment incidence rates by race for persons on welfare and persons not on welfare for each data set. We also compute the report rates for each and derive a measure of the reported maltreatment rates from these and ask: is it possible for there to be racial disparities in aggregated reported abuse rates when in fact there are no racial disparities in report rates or underlying abuse rates?

## **Results**

We begin with the final result and work our way backward. Table 4 shows the (derived) reported abuse rates for blacks and whites for all maltreatment cases in 1980, 1986 and 1993. These rates are the product of the underlying maltreatment rates per 1,000 found in the NIS data sets and the report rates for each year by race.

**Table 4: Derived Reported Maltreatment Rates  
1980, 1986, 1993**

	<b>Blacks</b>	<b>Whites</b>	<b>Percentage Difference</b>
<b>All Reporters</b>			
<b>NIS-1</b>	3.29	3.58	-8%
<b>NIS-2</b>	7.10	3.05	132%
<b>NIS-3</b>	9.55	5.42	76%
<b>Sampled Reporters</b>			
<b>NIS-1</b>	2.15	2.21	-3%
<b>NIS-2</b>	3.32	2.74	21%
<b>NIS-3</b>	6.84	4.01	71%

Source: Appendix Tables A.1 and A.2

Except for 1980, in each year black reported maltreatment rates are higher than the white reported maltreatment rates. The black and white rates are about the same in 1980. The derivations in Table 4 come by computing report rates for all sources obtained in NIS as well as report rates only for reporters sampled in the NIS design. Elsewhere we have argued that there is potential sample selection bias in measuring report rates when using NIS data without adjusting for the fact that some reports in the data set are tautologically equal to one (Ards, et al., 1998). Excluding the tautological reports reduces the report rates and thus reduces the reported maltreatment rates for both blacks and whites. Still, whether one looks at all reporters or just sampled reporters, the black and white reported maltreatment rates differ across years. The black reported maltreatment rate in 1993 was almost three times as high as it was in 1980, measured by either all reporters or sampled reporters. The white reported maltreatment rate also rose so that in 1993 it was almost twice that of 1980. But, since the black maltreatment rate rose faster than the white maltreatment rate, the nearly non-existent racial disparity found in 1980 became a wide disparity in 1993.

Could it be that in 1993 blacks appeared to be 1.71 times more likely to be maltreated than whites because whites are less likely than blacks to be on welfare and that welfare participants are

more likely to be maltreated? In other words, could the appearance of racial disproportionality in reported maltreatment be due to aggregation bias?

Table 5 breaks down the 1993 data to show rates for the subcategories AFDC recipients, non-AFDC recipients, abuse cases and neglect cases. In the appendix we detail how we produced the comparisons required to estimate maltreatment rates in the data set and to perform the significance tests. For simplicity we refer only to sampled reporters and, although data reported in most NIS documents combine whites and Hispanics, we refer only to non-Hispanic blacks and non-Hispanic whites. We produce in the appendix, however, our analysis using the NIS definitions of white as well.

**Table 5: Reported Maltreatment Rates, Report Rates and Maltreatment, 1993**

	Welfare Rates		Reported Maltreatment Rates (1) = (2) x (3)		Report Rates (2)		Maltreatment Rates (3)	
	Blacks	Whites	Blacks	Whites	Blacks	Whites	Blacks	Whites
Maltreatment								
AFDC	37%	7%	8.62	15.86	0.21	0.27	40.66	59.40
Non-AFDC	63%	93%	5.81	3.11	0.22	0.18	26.67	17.36
Total	100%	100%	6.84	4.01	0.22	0.20	31.82	20.36
								a,b
Abuse								
AFDC	37%	7%	3.82	9.66	0.28	0.32	13.83	29.91
Non-AFDC	63%	93%	4.23	2.76	0.36	0.31	11.87	8.97
Total	100%	100%	4.05	3.26	0.32	0.31	12.59	10.47
								c,d
Neglect								
AFDC	37%	7%	4.76	6.87	0.18	0.22	27.05	31.82
Non-AFDC	63%	93%	1.84	0.63	0.12	0.07	15.49	9.29
Total	100%	100%	2.90	1.06	0.15	0.10	19.74	10.90
					c,d	c		a,b

Key: Significance testing by statistical software used

	WesVar	Stata
p<.01	a	b
p<.05	c	d

Software used for analysis: WesVar v.4 for NIS-2 and NIS-3; Stata v.6 for NIS-1, NIS-2, and NIS 3

The first column of Table 5 displays the percentage by race of civilian, non-institutionalized children (under age 18) who receive welfare. Here we see that larger shares of black children receive welfare than do whites. The next column shows the derived measure of reported maltreatment rates for blacks and whites on welfare and not on welfare. The reported maltreatment rate for whites on

welfare is 15.86 per 1,000. The reported maltreatment rate for whites not on welfare is 3.11 per 1,000. Thus, the reported maltreatment rate for whites on welfare is five times as high as that for whites not on welfare. Among blacks, the difference is smaller. Blacks on welfare have reported maltreatment rates that are about one and half times that of blacks not on welfare. Overall, though, combining welfare and non-welfare recipients, it appears that blacks have a higher reported maltreatment rate than whites (6.84 compared to 4.01).

These reported maltreatment rates are derived by multiplying the report rate in the third column by the incidence rate of maltreatment computed in the fourth column. As it turns out, there are no statistically significant differences in the maltreatment report rates between blacks and whites or between welfare and non-welfare recipients. The maltreatment report rate for blacks on welfare is .21 and for blacks not on welfare, .22. The maltreatment report rate for whites on welfare is .27 and for whites not on welfare, .18, a difference that is not statistically significant due in part to the fact that so few whites are on welfare.

The last column of Table 5 shows the actual maltreatment incidence rates. There are large and statistically significant differences in the maltreatment rates of those on welfare and those not on welfare. Whites on welfare have substantially higher maltreatment rates than whites not on welfare: the rate is more than three times higher. The black maltreatment rate for those on welfare is 1.5 times as high as that for blacks not on welfare.

### **Discussion**

These statistics indicate that the appearance of higher reported maltreatment rates for blacks than for whites does not arise, in this instance, from higher report rates for blacks than for whites. Instead, it arises from higher actual maltreatment rates for welfare recipients (both white and black) than for those not on welfare, combined with the higher likelihood of blacks to be on welfare than whites. This is a problem of aggregation.

Another type of aggregation occurs in this example, providing both complexity and richness to our illustration. This aggregation arises because the abuse and neglect cases were combined. If we look just at abuse cases we find miniscule racial differences in reported abuse rates. We find no statistically significant differences in report rates between welfare and non-welfare recipients or between blacks and whites. But we do find statistically and significantly higher actual abuse rates among white welfare recipients as compared to white non-welfare recipients. Since there are few white welfare recipients this does not reverse the overall finding of no statistically significant racial differences in abuse.

Turning now to neglect, we find a different pattern. The report rate for white neglect victims is three times higher for those receiving welfare than for those not receiving welfare (.22 vs. .07). This difference is statistically significant, while, conversely, among black neglect victims, the difference between welfare and non-welfare report rates is not. There are no statistically significant differences in the report rates between blacks and whites, either. Actual neglect rates, moreover, are much higher for both black and white welfare recipients as compared to black or white non-welfare recipients. The result, then, is that black reported neglect rates are higher than white reported neglect rates overall—despite the fact that white welfare reported neglect rates are higher than black welfare reported neglect rates. The reason for the overall higher rates of neglect for blacks than for whites: the higher welfare participation rates for blacks than whites combined with the higher rate of neglect among welfare recipients than non-welfare recipients.

Aggregation creates the illusion of racial disproportionality when none may really exist.

## References

- Ards, S. (1992). Understanding patterns of child abuse. *Contemporary Policy Issues*, 10, 4, Oct., 39-50.
- Ards, S., & Harrell, A. (1993). Reporting of child maltreatment: A secondary analysis of the national incidence surveys. *Child Abuse & Neglect*, 17, 3, 337-344.
- Ards, S. Chung, C., & Myers, S.L., Jr. (February 1998). The effects of sample selection bias on racial differences in child abuse reporting” *Child Abuse & Neglect*, 22, 2, 103-115.
- Ards, S., Chung, C., & Myers, S. L., Jr. (August 1998). Reporting bias and child abuse. Roy Wilkins Center Working Paper.
- Blumstein, A. (1982). On the racial disproportionality of United States’ prison populations. *The Journal of Criminal Law and Criminology* 73, 3, 1259-1281.
- Cochran, W. G. (1977). *Sampling techniques, 3<sup>rd</sup> edition*. New York: John Wiley & Sons.
- Committee on Ways and Means, U.S. House of Representatives. 1989. *The 1989 green book: Background material and data on major programs within the jurisdiction of the Committee on Ways and Means*. Washington D.C.: U.S. G.P.O.
- Drake, B., & Zuravin, S. (April 1998). Bias in child maltreatment reporting: Revisiting the myth of classlessness. *American Journal of Orthopsychiatry* 68, 2, 295-304.
- Eltinge, J. L., & W. M. Scribney. (1996a). *Accounting for point-estimation bias in assessment of misspecification effects, confidence-set coverage rates and test sizes*. Unpublished manuscript. Department of Statistics. Texas A&M University.
- Eltinge, J. L. (1996b). Estimation of means, totals, ratios and proportions for survey data. *Stata Technical Bulletin* 31, 6-23.
- Giovannoni, J. M. & Becerra, R. M. (1979). *Defining child abuse*. New York: Free Press.
- Hindelang, M. J. (February 1978). Race and involvement in common law personal crimes. *American Sociological Review* 43,1, 93-109.
- Kovar, J.G., Rao, J. N. K., & Wu, C. F. J. (1988) Bootstrap and other method to measure errors in survey estimates. *Canadian Journal of Statistics* 16 (suppl.), 25-45. University of Toronto Press, New York, ON.
- Langan, P. A. (1985). Criminology—Racism on trial: New evidence to explain the racial composition of prisons in the United States. *The Journal of Criminal Law and Criminology* 76,3, 666-683.
- Myers, S. L., Jr. (1993). Measuring and detecting discrimination in the post-civil rights era. In J.H. Stanfield and R. M. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 172-197). Newbury Park, CA: Sage Publications.

- Myers, S. L., Jr. (1985). Statistical tests of discrimination in punishment. *Journal of Quantitative Criminology*, Summer.
- Macionis, J. J. (1987). *Sociology, fourth edition*. Englewood Cliffs, New Jersey: Prentice Hall, Inc.
- Morton, T. (December 1999). "Letter to the Editor" from Thomas Morton. *Child Abuse & Neglect*, 23, 12, 1209.
- Public Law 93-247. (1/31/74). S.1191. *Child Abuse Prevention and Treatment Act*.
- Public Law 98-457. (10/9/84). H.R. 1904. *Child Abuse Amendments of 1984, Family Violence Prevention and Services Act*.
- Public Law 100-294. (4/25/1988). H.R. 1900. *Child Abuse Prevention, Adoption, and Family Services Act of 1988*.
- Sedlak, A. J., Broadhurst, D., Shapiro, G., Kalton, G., Gostkel, H., Burker, J., & Brown, J. (1997, January). *Analysis report of the Third National Incidence Study of Child Abuse and Neglect*. Washington, D.C.: U.S. Department of Health and Human Services, Administration on Children Youth and Families, National Center on Child Abuse and Neglect.
- Sedlak, A. J., Hantman, I., & Schultz, D. (1997, February). *Third national incidence study of child abuse and neglect: Final report appendices*. Washington, D.C.: U.S. Department of Health and Human Services, Administration on Children Youth and Families, National Center on Child Abuse and Neglect.
- Sedlak, A. J., Hantman, I., & Schultz, D. (1997, May). *Third national incidence study of child abuse and neglect (NIS-3): Public use files manual*. Washington, D.C.: U.S. Department of Health and Human Services, Administration on Children Youth and Families, National Center on Child Abuse and Neglect.
- Stata reference manual release 6 volume 4: Su-Z*. (1999). College Station, TX: Stata Press.
- Stata technical bulletin 31*. (1996). College Station, TX: Stata Press.
- Turbett, J. P., & O'Toole, R. (1980). Physician's recognition of child abuse. Paper presented at the Annual Meeting of the American Sociological Association, New York. Quoted in Robert Hampton, *Violence against black children: Current knowledge and future needs*, "Violence in the Black family, Lexington Books, 1987.
- U.S. Bureau of Census. (1996). *Statistical abstract of the United States (116<sup>th</sup>)*, Table No. 22. Washington, DC: U.S. Bureau of the Census.
- U.S. Department of Health and Human Service Neglect. (1999). *Child maltreatment 1997: Reports from the states to the National Child Abuse and Neglect Data System*. Washington, DC: National Center on Child Abuse and Neglect (DHHS).

- U.S. Department of Health and Human Services. (1997). *Child maltreatment 1995: Reports from the states to the National Child Abuse and Neglect Data System*. Washington, DC: Administration on Children, Youth and Families (DHHS).
- U.S. Department of Health and Human Services. (1993). *Quarterly public assistance statistics*. Washington, D.C.: Social Security Administration, Office of Policy, Office of Research and Statistics.
- U.S. Department of Health and Human Services. (June, 1988). Table 3.9: Distribution of AFDC Children by Race, 1983-1996. *Aid to Families with Dependent Children: The baseline* [Online]. Available: [www.asp.hhs.gov/hsp/AFDC/afdcbase98exhib.htm](http://www.asp.hhs.gov/hsp/AFDC/afdcbase98exhib.htm)
- U.S. Department of Health and Human Services. (1987). *Quarterly public assistance statistics*. Washington, D.C.: Social Security Administration, Office of Policy, Office of Research and Statistics.
- U.S. Department of Health and Human Services. (September, 1981). *Study findings: National study of the incidence and severity of child abuse and neglect*. Washington, D.C.: National Center on Child Abuse and Neglect (DHHS).
- U.S. Department of Health and Human Services. (1980, Oct.; Nov.; Dec.). *Public assistance statistics*. Washington, D.C.: Social Security Administration, Office of Policy, Office of Research and Statistics.
- Wolter, K. M. (1985). *Introduction to variance estimation*. New York: Springer-Verlang.

## Appendix

### 1) Definitions of race

These analyses take the race of the children identified as experiencing maltreatment into account. Our primary concern was to identify differences between maltreatment incidence and reporting rates for black and white children. The NIS-1, NIS-2 and NIS-3 reports all include Hispanic children in the racial category described as “white” (see U. S. Dept. of Health and Human Services, 1981, Table 4-2 footnote 1, page 20). To avoid confusion, we have conducted our analyses using both the NIS definition, which we call “white & Hispanic” and our own definition, which we call “white/Non-Hispanic.” In some cases, excluding white Hispanic children from the comparison affects whether the difference in rates between white and black children is statistically significant. This occurs for Report Rates for Neglected Non-AFDC children in NIS-1 for All Reporters, Report Rates for Abused AFDC Children in NIS-2 for both All Reporters and Sampled Reporters, and Report Rates for Maltreated Children in NIS-3 for both All Reporters and Sampled Reporters. Another instance is found in the Incidence Rates for Neglected AFDC Children in NIS-1. When Hispanic children are excluded from the white category, the significance of the difference changes from  $< .01$  to  $< .05$ , due to the change in the number of cases.

### 2) Statistical packages and significance testing

The NIS-2 and NIS-3 data sets users guides strongly recommend WesVar (a statistical software package developed by Westat specifically for analysis of complex survey data) or a comparable statistical software package for analyzing the data. Analyses that do not allow for complex sampling designs often underestimate variance and thus overestimate the significance of results. We chose to replicate the analyses of NIS-2 and NIS-3 data using Stata, which also provides methods for calculating the variances for complex sampling designs. Our tables show results from both packages for these two data sets. WesVar uses Jackknife replication for estimating variance.

Stata uses a different technique (linearization). Our analyses show that in some cases WesVar and Stata produce different results when evaluating the significance of a difference in estimates. Interestingly, the differences are *not* consistent. Most often, Stata reports a difference to be more significant than WesVar, which would be expected if Stata's estimation of variance is less accurate than WesVar's. In one case (Incidence Rates for Black Neglected Children in NIS-3) Stata reports a significant result when WesVar does not.

The NIS-1 data were prepared before the development of WesVar and cannot be analyzed using that software package because they lack the required replicate weights. The analyses of NIS-1 data were done solely with Stata.

The following excerpts describe the WesVar and Stata methods and their applications to these data sets.

“Significance tests should be appropriate for the NIS-3 multi-stage sample design. The NIS-3 Public Use Data file is designed for use with WesVarPC (which will compute variances using the selected series of replicate weights in this data file).” *Third National Incidence Study of Child Abuse and Neglect (NIS-2): Public Use Files Manual* (Sedlak, Hantman & Schultz 1997, May, pp.1-18.)

“As described in the introductory chapter of this manual, the fact that the NIS uses a multi-stage sample design means that standard statistical packages do not appropriately compute the variances of study estimates (or the significance of differences). The NIS-3 Final Report and its appendices offer standard errors of estimates and the results of significance tests that were computed using a replication approach, the jackknife.\*” *Third National Incidence Study of Child Abuse and Neglect (NIS-2): Public Use Files Manual* (Sedlak, Hantman & Schultz 1997, May, pp. 2-88) \* is a footnote to Kovar, Rao & Wu's (1988) book, *Bootstrap and Other Method to Measure Errors in Survey Estimates*.

“The use of a stratified multi-stage cluster sample design, the adjustments for nonresponse and the adjustments for child duplication resulted in dependence among observations. A simple random sample formula for variance estimation is based on the assumption that observations are independently and identically (IID) distributed. Using such a formula for NIS-3 would underestimate the true sampling variance. In this survey, a replication method, jackknife, was used to estimate the sampling variance. This method provides unbiased estimates of variance in multi-stage cluster sample designs.” *Analysis Report of the Third National Incidence Study of Child Abuse and Neglect* (Sedlak, Broadhurst,, Shapiro, Kalton, Gostkel, Burkner and Brown, 1997, January, pp. 15-1)

In *Stata Reference Manual Release 6 Su-Z* (1999, pp. 67-68) the issue of survey estimation is discussed as follows:

“The current **svy** commands use the relatively simple variance estimators outlined below. See, for example, Cochran (1977) and Wolters (1985) for some methodological background on these variance estimators. In some cases, some authors prefer to use other variance estimators that, for example, account separately for variance components at different stages of sampling, use finite population corrections with some unequal-probability and multistage designs, and include other special design features.

In addition, the current **svy** commands use "linearization" based variance estimators for non-linear functions like sample ratios. Alternative variance estimators that use replication methods—for example, jackknifing balanced repeated replication—may be included in future versions."

In addition to Wolters (1985) cited by Westat, the other citation is to Cochran, W. G. (1977) *Sampling Techniques, 3rd ed.* New York: John Wiley & Sons. The chapter goes on to note that the procedures used in the **svy** commands:

"were developed in collaboration with John L. Eltinge, Department of Statistics, Texas A&M University." (p. 72)

Two publications by Eltinge are cited in the chapter references:

“Eltinge, J. L. and W. M. Scribney, 1996a *Accounting for point-estimation bias in assessment of misspecification effects, confidence-set coverage rates and test sizes.* Unpublished manuscript. Department of Statistics. Texas A & M University...1996b *svy2: “Estimation of means, totals, rations, and proportions for survey data.” Stata Technical Bulletin 31: 6-23. Reprinted in Stata Technical Bulletin Reprints, vol. 6, 213-235.*” (p. 72)

### 3) AFDC Child Population Estimation Methodology

In estimating the AFDC child population by race for each year, we had two primary considerations: 1) matching the AFDC population to the NIS sampling frame, and 2) using the best available data to estimate racial breakdowns of the AFDC child population. For 1986 (NIS-2) and 1993 (NIS-3), we used essentially the same approach. For each year, the aggregate AFDC child population came from the DHHS publications “Quarterly Public Assistance Statistics” (U.S. Department of Health and Human Services 1987; 1993). Because the NIS was conducted in the fall of each year, for 1986 we used the average AFDC population between October and December, while

for 1993 we used the AFDC population in September (the last month in 1993 for which published statistics were available). Given these figures, the next step was to break down the population by racial group. Figures for the racial breakdown of the AFDC population came from a June 1998 DHHS publication, “Aid to Families With Dependent Children: The Baseline” (U.S. Department of Health and Human Services, 1988, June). We applied the percentages for each racial group (White, Black, Other, and Hispanic—categories are mutually exclusive) to the total population to derive estimates of the AFDC child population by race.

One further step was necessary before using these figures in our incidence rate calculations: removing AFDC children in Puerto Rico, Guam, and the Virgin Islands from the totals (these areas were not in the NIS sampling frame, but are part of AFDC total population and racial percentage calculations). We assumed that all Puerto Rican AFDC children were Hispanic and subtracted these from the total number of Hispanic children derived in the previous step. While this may not be entirely accurate—there may be some Black or White non-Hispanic AFDC recipients in Puerto Rico—it was deemed accurate enough for our purposes, particularly since racial breakdowns of the Puerto Rican population are not available. For Guam and the Virgin Islands we assumed those AFDC child populations were distributed similarly to the national population. Again, while this may not be entirely accurate, it is accurate enough for our purposes (the number of AFDC children in Guam and the Virgin Islands is extremely small). After subtracting the AFDC child populations in Puerto Rico, Guam, and the Virgin Islands from the AFDC child populations by race, we arrived at the figures for AFDC child populations by race that we used in our incidence rate calculations.

For 1980, the basic method was the same, but we did not have direct estimates of the racial breakdown of the AFDC child population. The AFDC child population for 1980 was calculated as the average population between October and December (from monthly reports of *Public Assistance Statistics* for Oct. 1980, Nov. 1980, and Dec. 1980) (U.S. Department of Health and Human Services, 1980 Oct.; 1980, Nov.; 1980 Dec.). To arrive at these estimates, we explored six different methods of

estimating the racial breakdown of the 1980 AFDC child population by race. The first two methods used numbers of AFDC families by race in 1979 (available in the *Green Book* (1989), published by the House Ways and Means Committee), and applied different estimates of the average number of children per family for different races to the number of AFDC families by race. These methods provided estimates of the number of children in AFDC families by race, but relied upon strong assumptions about family size. The third method used the racial breakdown of poor children (from the Census) and applied these breakdowns to the AFDC child population.

The final three methods used the racial breakdowns of the AFDC child population from years where it was known. The first year for which AFDC children by race figures are available was 1983. We explored applying the 1983 breakdowns to the 1980 AFDC child population. The disadvantage of this approach is that in 1983, there is a substantial amount of missing data (roughly 10 percent of AFDC children are classified as race “unknown”). We also explored using the average racial breakdown between 1986 and 1993 (the other two years under consideration in this study), but it is unlikely that this number would reflect the racial breakdown of the AFDC child population in 1980.

Ultimately, we based our estimates on three figures: 1) the breakdown of AFDC families by race in 1979; 2) the breakdown of AFDC families by race in 1986; and 3) the breakdown of AFDC children by race in 1986 (1986 was chosen over 1983 because there was much less missing data). We assumed that the change in children by race from 1979 to 1986 mirrored the change in families by race from 1979 to 1986. Thus, we calculated a factor for each racial group equal to the (percent of families of that race, 1979) divided by the (percent of families of that race, 1986). We then applied this factor to the percent of *children* of that race in 1986, which resulted in an estimated percent of children of that race in 1979. After normalizing these estimates so they summed to 100 percent, we applied these normalized figures for racial breakdowns to the total AFDC child population for 1980 to derive estimates of the child population of each race. These totals were adjusted to remove the

territorial and Puerto Rican AFDC populations according to the method described above for 1986 and 1993.

NOTES

1. All rates calculated by us from NIS data. Overall rates agree with figures in NIS published final reports.
2. "White/Non-Hispanic" is this paper's definition of "white." The NIS definition of "white," used in the NIS final reports, includes both whites and Hispanics (of any race).
3. "Non-AFDC" incidence numbers above include children in NIS datasets whose AFDC status was "No" or "Unknown".

NIS-1 (1980)	ALL REPORTERS						SAMPLED REPORTERS					
	All Races	Black	White/Non-Hisp (note 2)	Sig. Of Diff. W-B	White & Hisp (note 2)	Sig. Of Diff. W&H-B	All Races	Black	White/Non-Hisp (note 2)	Sig. Of Diff. W-B	White & Hisp (note 2)	Sig. Of Diff. W&H-B
<b>Maltreatment</b>												
AFDC	0.43	0.35	0.47		0.45		0.27	0.22	0.29		0.29	
Non-AFDC	0.29	0.24	0.29		0.30		0.19	0.17	0.18		0.19	
Total	0.33	0.28	0.33		0.33		0.21	0.19	0.21		0.21	
Sig of Diff, AFDC-non	d		d		d							
<b>Abuse</b>												
AFDC	0.47	0.65	0.45		0.44		0.30	0.47	0.28		0.27	
Non-AFDC	0.39	0.53	0.37		0.38		0.26	0.42	0.24		0.25	
Total	0.41	0.57	0.39	d	0.39	d	0.27	0.43	0.25	d	0.25	d
Sig of Diff, AFDC-non							d					
<b>Neglect</b>												
AFDC	0.36	0.23	0.42		0.41		0.22	0.14	0.25		0.25	
Non-AFDC	0.18	0.08	0.20		0.20	d	0.11	0.05	0.12		0.12	
Total	0.43	0.14	0.26		0.26		0.14	0.08	0.15		0.15	
Sig of Diff, AFDC-non	b	d	b		b		d				d	

NIS-2 (1986)	ALL REPORTERS						SAMPLED REPORTERS					
	All Races	Black	White/Non-Hisp (note 2)	Sig. Of Diff. W-B	White & Hisp (note 2)	Sig. Of Diff. W&H-B	All Races	Black	White/Non-Hisp (note 2)	Sig. Of Diff. W-B	White & Hisp (note 2)	Sig. Of Diff. W&H-B
<b>Maltreatment</b>												
AFDC	0.32	0.27	0.38		0.33		0.19	0.15	0.26		0.22	
Non-AFDC	0.30	0.23	0.32		0.32		0.20	0.14	0.21		0.21	
Total	0.31	0.24	0.33		0.32		0.19	0.14	0.22		0.21	
Sig of Diff, AFDC-non												
<b>Abuse</b>												
AFDC	0.44	0.27	0.59	c, d	0.52		0.28	0.17	0.46	c, d	0.38	
Non-AFDC	0.42	0.34	0.44		0.43		0.30	0.21	0.32		0.32	
Total	0.42	0.31	0.46		0.44		0.29	0.20	0.34		0.33	
Sig of Diff, AFDC-non												
<b>Neglect</b>												
AFDC	0.22	0.26	0.19		0.17		0.11	0.13	0.12		0.10	
Non-AFDC	0.15	0.15	0.15		0.15		0.07	0.09	0.08		0.07	
Total	0.17	0.19	0.16		0.15		0.08	0.10	0.08		0.08	
Sig of Diff, AFDC-non												

NIS-3 (1993)	ALL REPORTERS						SAMPLED REPORTERS					
	All Races	Black	White/Non-Hisp (note 2)	Sig. Of Diff. W-B	White & Hisp (note 2)	Sig. Of Diff. W&H-B	All Races	Black	White/Non-Hisp (note 2)	Sig. Of Diff. W-B	White & Hisp (note 2)	Sig. Of Diff. W&H-B
<b>Maltreatment</b>												
AFDC	0.33	0.30	0.35		0.35		0.25	0.21	0.27		0.27	
Non-AFDC	0.28	0.31	0.24		0.24		0.20	0.22	0.18		0.18	
Total	0.28	0.30	0.27		0.27		0.21	0.22	0.20		0.20	
Sig of Diff, AFDC-non			d		c, d						c, d	
<b>Abuse</b>												
AFDC	0.40	0.32	0.40		0.42		0.34	0.28	0.32		0.35	
Non-AFDC	0.40	0.45	0.38		0.38		0.32	0.36	0.31		0.31	
Total	0.40	0.40	0.39		0.39		0.33	0.32	0.31		0.32	
Sig of Diff, AFDC-non												
<b>Neglect</b>												
AFDC	0.28	0.28	0.31		0.29		0.20	0.18	0.22		0.21	
Non-AFDC	0.14	0.21	0.12		0.12		0.09	0.12	0.07		0.08	
Total	0.18	0.24	0.16		0.16		0.12	0.15	0.10		0.11	

Key: Significance testing by statistical software used

	WesVar	Stata
p<.01	a	b
p<.05	c	d

Software used for analysis: WesVar v4 for NIS-2 and NIS-3; Stata v. 6.0 for NIS-1, NIS-2, and NIS-3.