

Assessing Parenting Behaviors across Racial Groups: Implica- tions for the Child Welfare System

Lawrence M. Berger
University of Wisconsin–Madison

Marla McDaniel
Columbia University

Christina Paxson
Princeton University

Black families are disproportionately represented in the child welfare system. This may in part result from racial bias in judgments made by those who report and investigate child maltreatment. However, little is known about how race influences judgments about parenting. This article relies on data from a population-based survey to examine whether the race of interviewers, relative to the race of families they interview, influences parenting assessments. It reports evidence of racial bias in some measures of interviewer-assessed parenting behaviors. Racial bias is more pronounced for measures that require subjective assessments on the part of interviewers.

In 2002, approximately 12.3 of every 1,000 U.S. children were determined to have been abused or neglected (U.S. Department of Health and Human Services 2002). However, the incidence of child maltreatment is not evenly distributed. Official child maltreatment rates are considerably higher for black families than for white families. A large body of research indicates that, compared with white families, black families are disproportionately reported and investigated for child abuse and neglect. So, too, reports are disproportionately substantiated

Social Service Review (December 2005).
© 2005 by The University of Chicago. All rights reserved.
0037-7961/2005/7904-0004\$10.00

(Courtney et al. 1996; Garland et al. 1998; Fluke et al. 2003). Approximately 15 percent of the population of U.S. children is black, yet black children account for 25 percent of substantiated child maltreatment victims. In contrast, about 65 percent of U.S. children are white, but only 51 percent of substantiated victims are white (U.S. Bureau of the Census 2001, table 19; U.S. Department of Health and Human Services 2002). Official maltreatment rates for these groups have remained relatively stable in recent years.

There is an ongoing debate over the sources of racial disparity in official maltreatment rates (Levine et al. 1996; Roberts 2002; Ards et al. 2003; Hines et al. 2004). One explanation is that racial disparities in child maltreatment may be driven by differences in poverty rates across racial groups. Black families are more likely than white families to live in poverty, and poverty is highly correlated with child maltreatment.¹ A second explanation is that parenting behaviors may differ across racial groups and that such differences may result from cultural or other factors correlated with race. A third explanation is that there may be racial bias in the child protective services (CPS) systems. The term "racial bias" has been broadly used to represent several conceptually distinct reasons for elevated rates of maltreatment among blacks (Drake and Zuravin 1998). However, in most cases, the term connotes a racial double standard such that, given identical behaviors, black parents are more likely than white parents to face charges of maltreatment and that charges against them are more likely to be substantiated.

Despite this long-standing debate, relatively little theoretical or empirical research examines racial bias in child maltreatment reporting. While the existing literature offers a useful categorization of the routes through which racial bias might affect the multiple judgments that may lead to a substantiated case of maltreatment, there is little discussion of how individuals form judgments about the parenting behaviors of others and whether these judgments are biased. Furthermore, empirical investigations of racial bias in reporting are complicated by the fact that maltreatment reports come from a large and heterogeneous population. Child protective service systems rely on an army of volunteers, including doctors, social workers, teachers, neighbors, relatives, and strangers, to make reports of maltreatment.² In 1999, 55 percent of screened-in referrals came from professionals, and 45 percent came from nonprofessionals (U.S. Department of Health and Human Services 2002).³ Empirical studies of racial bias, discussed in more detail below, typically focus on the behaviors of narrowly defined professional groups that have experience working with families and children. It is possible that many members of these groups have been trained to assess child maltreatment. However, we know of no existing studies that examine whether individuals without such expertise (who comprise nearly half of all reporters) make racially biased judgments.

This article focuses on the process by which individuals make judgments about the parenting behaviors of others and explores whether these judgments are influenced by race. Data from a population-based survey of parents with preschool age children are used to investigate how race is associated with self-reported and interviewer-assessed parenting behaviors. The work examines whether there are racial differences in parenting behaviors across each type of measure, and whether racial differences in socioeconomic status and maternal characteristics can explain any observed differences. Results from these analyses show that there are racial differences in both self-reported and interview-assessed parenting measures, and these differences cannot be fully explained by other factors. The work then investigates whether the race of the interviewer, in and of itself and in combination with the race of the parent, influences assessments of parenting. The findings provide evidence of racial bias in some (but not all) measures of parenting behaviors.

Background

Definitions of Racial Bias in Maltreatment

A substantiated case of child maltreatment is the result of (at least) a three-step process. First, a family must come into contact with someone who will potentially report an alleged incident to CPS. Second, the potential reporter must decide whether to file a report. Third, if a report is made and investigated, CPS must decide whether the allegation of abuse should be substantiated.⁴ After substantiation, family court judges and caseworkers make decisions about whether the child should be removed from the home and, if so, how long the child will stay in out-of-home care.

The literature on racial bias (and, more generally, on class bias) in child maltreatment has defined bias according to the specific points at which it occurs along this process (Drake and Zuravin 1998). “Exposure bias” refers to the idea that contact with potential reporters may vary systematically across racial, ethnic, and income groups. For example, black children may be more likely than white children to come into contact with mandated reporters, because of their greater rates of involvement with the welfare system and other government programs that serve the poor. “Reporting bias” refers to the idea that members of some racial, ethnic, or economic groups are more likely than others to be reported, even if their observed behaviors are identical. “Substantiation bias” means that reports of maltreatment against members of some groups are more likely to be substantiated than reports against members of other groups, given identical information in the report or uncovered during the course of the investigation. Bias may also creep into case disposition decisions, including whether to remove a child from a home

and the length of time children remain in out-of-home care (Drake and Zuravin 1998; Ards et al. 2001; Harris and Courtney 2003; Needell, Brookhart, and Lee 2003; Wulczyn 2003).

Sources of Racial Bias

Reporting and substantiation bias differ from exposure bias in that they arise when potential reporters, caseworkers, or judges make decisions that differ by the race of the family, holding observed parenting behavior fixed. There are several reasons why racially biased reporting and substantiation decisions may be made. One is that potential reporters, caseworkers, or judges may believe that members of one racial group are less likely to challenge their decisions than members of another group. This reason, although possibly important in practice, is not relevant to this study, since the interviewers studied here are called upon to make judgments about parenting but not to make reporting decisions.

Another reason for racially biased decisions is that there may be racial bias in the judgments about parenting. This reason is relevant to this study because theories of statistical discrimination provide a useful framework for understanding racial biases in judgments.⁵ These theories rely on the idea that observers may have incomplete information; in this case, observers may have incomplete information about what has actually happened to a child in a specific situation. They may therefore rely on their beliefs about how parents of a particular race typically behave in order to draw inferences about the truth. For example, a potential reporter may observe that a child has an injury that could be the result of maltreatment but could also be due to an accident. A reporter who believes (correctly or not) that black families are more likely to use physical discipline strategies may then be more likely to infer that a black child has been maltreated than has a white child with an identical injury.

Several points about this theory require emphasis. First, reporters' beliefs about the typical behaviors of members of different racial groups could be based on the true distribution of behaviors in the population. If so, reporters of the same race as the family being judged would be no less likely to display racial bias than reporters of a different race. However, reporters' beliefs may also be based on stereotypes that need not match reality. In this case, it is reasonable to expect that bias will be most common when families are being judged by observers from another racial group. Second, incomplete information could lead to more subtle forms of bias. For example, potential reporters, especially those of different racial or ethnic backgrounds than the families being judged, may be ill-equipped to interpret cultural cues or understand the meanings of behaviors they observe. If so, racial bias should again

be most common when families are being judged by members of other racial or ethnic groups.

The empirical work that follows examines two implications of racial bias in judgments. First, judgments about situations that are ambiguous, and therefore require subjective interpretation on the part of observers, are more likely to result in racially biased judgments than situations that are clear-cut. Second, provided that observers of different races have varying beliefs about the typical behaviors of families of different races, the amount of racial bias will depend on the race of the observer relative to the race of the family being assessed. A comparison of the ratings that white observers give to black families relative to white families, with the ratings that black observers give to black families relative to white families, provides information on whether there is racial bias in judgments.

Although both of these empirical implications are consistent with racial bias in judgments, it should be noted that the second implication—that judgments will depend on the race of the observer relative to the race of the family—could possibly stem from a different source. Specifically, families may tend to behave differently when being observed by someone from a different racial group than when being observed by someone from their own racial group. The members of any family are unlikely to feel completely at ease while being questioned by a social worker, physician, or (in our case) interviewer. The level of anxiety felt by parents and children may affect how they communicate with the observer and how they interact with each other. These effects could be heightened when families are assessed by someone of a different racial group, possibly increasing the likelihood that black parents will be assessed negatively by white observers, and vice versa. These kinds of racial interaction effects have been documented in other contexts. For example, Hyoshin Kim, Mazli Baydar, and April Greek (2002) use the National Longitudinal Survey of Youth to show that black children score lower on cognitive tests administered by white interviewers than they do on tests given by black interviewers. It is not clear whether this is due to greater anxiety on the part of children when faced with a white interviewer, to less encouragement from white interviewers, or to lower expectations on the part of white interviewers when testing black children. Nevertheless, this evidence suggests that behavior (in this case, test performance) is affected by racial interactions. The empirical methods described below cannot distinguish between racial bias in judgments and changes in behaviors produced by racial interactions.

Evidence on Racial Bias in Judgments

Reporting and substantiation decisions rely on judgments by potential reporters, caseworkers, or judges about the quality of care children have received. A small but growing empirical literature uses two distinct ap-

proaches to examine whether such judgments are influenced by race. The first approach is based on vignettes. In this approach, individuals, such as medical professionals, teachers, and social workers, are given a set of scenarios and asked to indicate whether each described scene constitutes maltreatment (Zellman 1992; Pakieser, Starr, and LeBaugh 1998; Egu and Weiss 2003). To examine whether judgments are influenced by the race of the child, the researcher changes the vignettes by manipulating the child's race without altering any other element of the scene. Some vignette studies also examine whether the race of the assessor influences judgments. The second approach is based on the study (through chart or case reviews) of actual reporting and substantiation decisions made by professionals who come into contact with children and families (Hampton and Newberger 1985; Lane et al. 2002). In this approach, researchers assess whether similar types of cases are subject to different reporting (or substantiation) decisions based on the race of the child or family. A clear advantage of vignettes over case review studies is that the researcher carefully controls all elements of the scenario. However, it is possible that case review studies provide better information about actual behavior.

These two types of studies tend to yield different results. In general, vignette studies find little evidence of racial bias in judgments. For example, one study of teachers examines how maltreatment severity, the race of the teacher, and the race of the child affect reporting decisions (Egu and Weiss 2003). Although case severity affects teachers' (hypothetical) decisions to report a case, the races of the teacher and child do not significantly influence reporting decisions. Another vignette study surveys pediatricians, psychologists, and social workers, finding that the child's race does not influence reports (Zellman 1992). However, case severity, the child's age, and (to a lesser extent) socioeconomic status do affect these decisions. A vignette study of nurses also concludes that the race of the child does not influence outcomes (Pakieser et al. 1998).

In contrast, evidence from case review studies tends to support the hypothesis of racial bias in reporting. Two studies evaluate reporting decisions among hospital staff and find evidence that physicians are more likely to report black children than white children who are seen for similar physical injuries (Hampton and Newberger 1985; Lane et al. 2002). The latter of these studies finds that white toddlers with skull or long-bone fractures are less likely to be reported to CPS than minority toddlers, even if the specific nature of the injury is fixed. Research also indicates that black and low-income infants are more likely than infants from white and more affluent families to be tested for drug exposure at birth. This may subsequently result in a CPS report, despite similar rates of prenatal substance use across racial and socioeconomic groups (Chasnoff, Landress, and Barrett 1990; Sagatun-Edwards and Saylor 2000).

As noted above, case review studies have the advantage that, unlike

the vignette studies, they examine actual rather than hypothetical reporting decisions. However, case review studies have several drawbacks. First, they rely almost exclusively on reports by medical professionals, for whom objective chart data are more readily available. It is not known whether other types of reporters make similar judgments. Second, it is difficult to determine whether decisions that appear to be racially biased may have been influenced by other relevant information not included in the case record. For instance, a physician who treats a child for a fracture may have had an opportunity to question family members about the incident, to observe their demeanors, and to note their interactions with each other. He or she may also have access to information about the child's medical history that is not reflected in the records available to researchers. This information can result in what appears to be racial bias if it is correlated with race.

The empirical analyses presented below use survey data to investigate whether there is racial bias in interviewer assessments of parenting behaviors. This follows in the tradition of the case review studies, rather than vignette studies, examining actual judgments made by interviewers about the parenting behaviors they observe. However, the use of survey data presents several advantages that are not available in most case review studies. First, the survey data contain information on parents' reports of their own behaviors. Racial differences in these measures cannot be easily attributed to racial bias on the part of the interviewer (although it is possible that the race of the interviewer influences the answers provided). Second, the data contain information on a large sample of both black and white interviewers, most of whom interviewed both black and white parents. This allows an examination of whether the judgments of individual interviewers vary across families whose races do and do not differ from their own.

Data and Measures

Sample

The data are drawn from an in-home module of the Fragile Families and Child Wellbeing Study (FFCWS). This is a longitudinal birth cohort study that began in 1998 with a baseline sample of approximately 4,800 births in 20 U.S. cities (for a complete description of the sample and design, see Reichman et al. 2001). The survey contains an oversample of nonmarital births. As a consequence, children in this sample differ from those in a nationally representative sample.⁶ They are more likely to be poor, to have absent fathers, and to have mothers with low levels of education. Previous research indicates that children with these characteristics are more likely than others to be reported to their states' child protective service agencies (Gil 1970; Hampton and Newberger

1985; Zellman 1992; Lindsey 1994). The sample is also racially diverse: 47 percent of the mothers originally sampled identified themselves as non-Hispanic black, and 27 percent identified as Hispanic.

Families were surveyed three times: in person at the time of the child's birth, by telephone when the child was 12 months, and again by telephone at 36 months of age. Following the 36-month interview, families were asked to participate in an in-home assessment. The in-home module uses both a questionnaire and a set of interviewer observations. It assesses multiple domains of parenting, the child's home environment, and mother-child interactions.⁷ There was no attempt to match black and white respondents with same-race interviewers. The analyses that follow are based on a subsample of 1,417 children whose mothers were either classified as black or white, who participated in the in-home module, and for whom the race of the interviewer was known.⁸ Hispanic respondents are excluded from the analyses because, for language reasons, they were often purposefully assigned to Hispanic interviewers. As discussed below, the tests for racial bias require essentially random assignment of interviewers to families.

Measures of Parenting, Maternal, and Child Characteristics

Ten measures of parenting, maternal characteristics, and child characteristics are summarized below. Complete details and descriptive statistics on each scale are presented in the appendix. There are three self-reported and two interviewer-assessed parenting measures. The self-reported parenting measures consist of three subscales drawn from the Parent-Child Conflict Tactics Scales (CTS), which measure parents' use of different discipline strategies (Straus and Gelles 1990; Straus et al. 1998). The specific subscales are (1) use of nonviolent discipline strategies, (2) use of psychological aggression, and (3) use of physical assault.⁹ The CTS collect information on the parent's self-reported use of these strategies during the 12 months prior to the interview. Scores used here measure the annual incidence of each type of behavior. Use of nonviolent discipline strategies is reverse coded, so that higher scores correspond with less use of nonviolent strategies, such as putting children in time-outs or taking away privileges.

The two interviewer-assessed parenting measures are drawn from subscales of the Home Observation for Measurement of the Environment (HOME; see Caldwell and Bradley 1984; Bradley 1993).¹⁰ These particular subscales measure harshness and lack of warmth.¹¹ The scale for harshness includes five items that measure the mother's behavior toward the child during the visit. These document whether or not the mother shouted at, expressed annoyance with, expressed hostility toward, slapped, spanked, scolded, criticized, and excessively interfered with or restricted the child during the visit. The lack of warmth scale consists

of six interviewer-observed items that measure whether the mother spontaneously vocalized to the child, responded verbally to the child's vocalizations, introduced the child to a person or object, praised the child, conveyed positive feelings to the child, and caressed or kissed the child during the interview.¹² Respondents are given one point for each item that is observed during the interview. Items are then summed within each subscale to obtain a total subscale score. The subscales are coded such that higher scores represent greater degrees of harshness and less warmth.

In addition to these five measures of parenting, there are also five interviewer-assessed measures of maternal and child characteristics and behaviors. These measures are (1) lack of maternal verbal and social skills, (2) lack of maternal attention and understanding, (3) maternal hostility and suspicion, (4) problems with the child's appearance, and (5) problems with the child's behavior. The measure of lack of verbal and social skills is made up of three items that focus on the mother's interactions with the interviewer. It includes whether she exhibited audible and distinct speech, initiated verbal exchanges, and conversed freely. The measure of lack of maternal attention and understanding is based on items that rated the mother's attention to the interviewer, understanding of the questions, ability to articulate answers, and level of cooperation. The maternal hostility and suspicion measure includes items that assess whether the mother appeared suspicious, uncommunicative, anxious, nervous, or hostile, as well as whether she appeared to be on drugs. The measure of problems with the child's appearance is based on 14 items relating to the condition of the child's clothing and his or her degree of cleanliness. Finally, the measure of problems with the child's behavior is based on five items that assess the child's behavior during the interview. These include the child's persistence, cooperation, and displays of emotion.

Each scale is transformed into discrete measures indicating whether the score falls above a threshold value at which higher values correspond with more problematic behavior. Because none of the scales has standard thresholds to indicate problematic behavior, two discrete measures are constructed for each scale. One is an indicator for whether the score falls above the seventy-fifth percentile for the overall sample; the other is an indicator for whether the score exceeds the ninetieth percentile. Descriptive statistics for these measures are shown in table 1. Note that some of the measures are heavily skewed, with the majority of parents receiving perfect (low) scores. For one measure (maternal suspicion and hostility), more than 75 percent of parents have a score of zero, so only the determinants of falling above the ninetieth percentile cut point can be examined. This table also presents *t*-statistics for the hypothesis that blacks and whites are equally likely to have high scores. This hypothesis is strongly rejected for all scales except those measuring

Table 1

DESCRIPTIVE STATISTICS FOR PARENTING AND MATERNAL AND CHILD BEHAVIORS AND CHARACTERISTICS

	INDICATOR: SCORE AT OR ABOVE 75TH PERCENTILE				INDICATOR: SCORE AT OR ABOVE 90TH PERCENTILE			
	75th Percentile Cut Point	White	Black	t-test: White = Black	90th Percentile Cut Point	White	Black	t-test: White = Black
Maternal-assessed parenting measures:								
Lack of nonviolent discipline	61	.09	.27	6.80	76	.03	.10	3.91
Psychological aggression	48	.19	.25	2.32	54	.08	.11	1.53
Physical assault	32	.20	.26	2.34	54	.04	.09	2.61
Interviewer-assessed parenting measures:								
Harshness	1	.09	.28	7.59	2	.03	.11	4.37
Lack of warmth	1	.08	.17	4.23	3	.01	.09	5.07
Interviewer-assessed maternal and child behaviors and characteristics:								
Lack of maternal verbal and social skills	0	.05	.15	4.87	1	.03	.07	2.86
Lack of maternal attention and understanding	3	.12	.28	5.98	6	.02	.08	3.70
Maternal hostility and suspicion	NA	NA	NA	NA	0	.03	.06	2.36
Problems with child's appearance	0	.16	.26	3.75	4	.03	.07	2.47
Problems with child's behavior	7	.20	.26	2.19	12	.06	.09	1.91
Observations		337	1,080			337	1,080	

NOTE.—NA = not applicable. The 75th percentile cut point columns provide the values of the scales below which at least 75% of the total sample falls. The 90th percentile cut point columns provide the values of the scales below which at least 90% of the total sample falls. The columns titled "White" and "Black" indicate the fraction of whites and blacks who score at or above the cut point. The columns marked t-test provide results of a t-test of whether the fraction of whites and blacks who fall above the cut point are equal. Appendix table A1 shows more complete information on the frequency distributions of the scales and χ^2 tests of the null hypothesis that the distributions for blacks and whites are identical.

the use of psychological aggression and problems with the child's behavior (at the ninetieth percentile cut point). In general, blacks have higher scores than do whites, and this indicates more problematic scores.¹³

Household and Maternal Characteristics

Racial differences in parenting behaviors may result from differences across racial groups in socioeconomic characteristics, such as income, education, and family structure, as well as from differences in maternal characteristics that are correlated with race. The analyses control for measures of family socioeconomic status and maternal characteristics. These measures are drawn from the core telephone survey. Measures of the family's socioeconomic status include the logarithm of family income adjusted for family size; the numbers of children and adults in the household; indicators for whether the mother has less than a high school degree, a high school degree (i.e., high school diploma or general equivalency diploma), or more than a high school degree; an indicator for whether the mother is single (defined as neither married nor cohabiting with a partner); indicators for the presence and employment status of the father or other man in the household; and an indicator for whether the mother worked in the week before the survey.¹⁴ Maternal characteristics include an eight-point maternal depression scale drawn from the composite international diagnostic interview—short form (Kessler et al. 1998) and a set of indicator variables for whether the mother reported that she engaged in various risky prenatal behaviors. These variables include whether she smoked cigarettes, drank alcohol, or used drugs during her pregnancy with the child.

Table 2 shows means for a set of observable parental characteristics for blacks and whites. It provides evidence that differences in household and maternal characteristics are large. Test statistics presented in the final column of the table indicate that the two racial groups significantly differ on 10 of the 15 household and maternal characteristics. In comparison with white families, on average, black families have considerably lower incomes, more children, and fewer adults in their households. Black mothers also have less education and are more likely to be single or to report living with a man who is not the focal child's father and is also not working. Consistent with findings from other literature, black mothers are less likely than white mothers to report that they smoke and drink during pregnancy.

Interviewer Characteristics

The 1,417 respondents were interviewed by 91 interviewers, with an average number of 15.6 interviews per interviewer (interviews per interviewer ranged from one to 52). Forty-two percent of the interviewers

Table 2

AVERAGE HOUSEHOLD AND MATERNAL CHARACTERISTICS

	White	Black	t-test
Sociodemographic characteristics:			
Ln (income-to-adult equivalent)	9.46 (.95)	8.49 (1.04)	15.10
Number of children	1.96 (1.15)	2.43 (1.42)	5.48
Number of adults	1.99 (.60)	1.84 (.86)	3.04
Mother works	.590	.563	.86
Mother has high school diploma or GED	.258	.363	3.56
Mother has more than high school diploma or GED	.564	.284	9.69
Child's father at home and working	.638	.289	12.15
Child's father at home and not working	.074	.094	1.13
Other man at home and working	.068	.082	.83
Other man at home and not working	.006	.030	2.50
Neither child's father nor other man at home	.214	.505	9.70
Maternal depression and risky behaviors:			
Maternal depression score	1.36 (2.47)	1.55 (2.55)	1.24
Smoked while pregnant	.297	.210	3.30
Drug use while pregnant	.047	.068	1.33
Alcohol use while pregnant	.157	.106	2.57
Observations	337	1,080	

NOTE.—Ln = natural logarithm; GED = general equivalency diploma. The columns titled "White" and "Black" present mean (and standard deviation) statistics. The t-test column provides results of a t-test that the mean value for whites and blacks is equal.

were black, and 58 percent were white.¹⁵ Seventy-nine percent of black interviewers and 66 percent of white interviewers interviewed both black and white mothers.

For the analyses that follow, it is important that the assignment of interviewers to families was essentially random; correlation between the race of the interviewer and unobservable determinants of parenting behavior would result in biased parameter estimates. Although the randomization of interviewers by race was not an explicit part of the survey design, there was no systematic attempt to match households to interviewers.¹⁶ However, because cities with higher fractions of blacks in their populations tended to have higher fractions of both black interviewers and black respondents, the race of interviewers and respondents is correlated.

In table 3, an indicator for whether a black interviewer was assigned to the family is regressed on an indicator for whether the mother is black, as well as on the complete set of sociodemographic and maternal controls. The first column shows results from a probit regression that does not include city dummies. The results show that, without control-

Table 3

DETERMINANTS OF ASSIGNMENT OF INTERVIEWERS TO RESPONDENTS

	Probit, No City Dummies	OLS, No City Dummies	OLS, City Dummies
Mother is black	.136*** (.034)	.135*** (.035)	.012 (.026)
<i>F</i> -test (joint insignificance of socio- demographic and maternal controls) ^a	.319	.317	.538
<i>F</i> -test (joint insignificance of city dummies) ^a			.000

NOTE.—Dependent variable: indicator that interviewer is black. OLS = ordinary least squares. 1,417 observations. All models include the maternal and household characteristics listed in table 2. Marginal effects are presented. Standard errors are in parentheses.

^a Denotes *p* value.

*** *p* < .001.

ling for city, black mothers are 13.6 percentage points more likely than white mothers to be assigned a black interviewer. (The sociodemographic and maternal controls are jointly insignificant.) Similar results are obtained from an ordinary least squares (OLS) regression in column 2. However, when city dummies are included (col. 3), black mothers are found to be only 1.2 percentage points more likely than whites to be assigned to a black interviewer. That difference is not statistically significant. Thus, it appears that racial differences in assignment of interviewers to mothers are completely accounted for by the racial composition of the cities under study. For this reason, a complete set of city indicators is included in all models.

Empirical Analyses

Methods

The first step in this analysis is to examine whether there are racial differences in the maternal- and interviewer-assessed outcome measures, as well as the extent to which such differences can be explained by other observable maternal and child characteristics. A series of OLS models are estimated for the probability that a family scores in the seventy-fifth and ninetieth percentiles for each of the parenting outcomes.¹⁷ The OLS regressions are of the form

$$y_R = \beta_0 + \beta_{RB}I(R \text{ Black}) + \mathbf{X}_1\alpha_1 + \mathbf{X}_2\alpha_2 + \mathbf{X}_3\alpha_3 + \varepsilon, \quad (1)$$

where y_R is an outcome for respondent R , $I(R \text{ Black})$ is an indicator that the mother is black, \mathbf{X}_1 is a vector of sociodemographic characteristics, \mathbf{X}_2 is a vector of maternal depression and behaviors, \mathbf{X}_3 is a set of city dummy variables, and ε is a disturbance term. Three variants of this

model are estimated for each outcome. The first (model 1) includes only the maternal race variables and city fixed effects as predictors. The second (model 2) adds sociodemographic characteristics. The third (model 3) adds the maternal depression and risky behavior variables. Given the high degree of correlation among the control variables, the models cannot be expected to estimate the effects of any of the controls precisely (and coefficients for individual control variables are not reported on the tables). Instead, the focus is on tests of joint significance for each set of controls and on whether the addition of the controls alters estimates of the race parameter β_{RB} .

The second step in the analysis is to investigate whether the race of the interviewer, alone and in interaction with the mother's race, influences outcomes. Here, OLS regressions of the same form as those presented in the first equation are estimated, but with the addition of an indicator for whether the interviewer is black:

$$y_{RI} = \beta_0 + \beta_{RB}I(R \text{ Black}) + \beta_{IB}I(I \text{ Black}) + \mathbf{X}\alpha + \varepsilon, \quad (2)$$

where y_{RI} is an outcome for respondent R interviewed by interviewer I , $I(I \text{ Black})$ is an indicator that the interviewer is black, and \mathbf{X} denotes all other controls (sociodemographic characteristics, maternal characteristics, and city indicators). This specification allows for an examination of whether black and white interviewers systematically rate respondents differently. It does not, however, provide information on racial bias in assessments.

The third step is to estimate these models with a full set of interactions between the interviewer's race and the respondent's race. The omitted category is white respondent and white interviewer. Specifically, equations of the following form are estimated:

$$y_{RI} = \beta_0 + \beta_{RB,IW}I(R \text{ Black}, I \text{ White}) + \beta_{RW,IB}I(R \text{ White}, I \text{ Black}) \\ + \beta_{RB,IB}I(R \text{ Black}, I \text{ Black}) + \mathbf{X}\alpha + \varepsilon. \quad (3)$$

Estimates of equation (3) provide the basis for our tests of racial bias. These estimates examine whether the difference in assessments of black and white respondents by white interviewers (i.e., $\beta_{RB,IW} - \beta_{RW,IW}$, where $\beta_{RW,IW}$ has been normalized to zero) is equal to the difference in assessments of black and white respondents by black interviewers (i.e., $\beta_{RB,IB} - \beta_{RW,IB}$). Note that although a rejection of the hypothesis that $\beta_{RB,IW} - \beta_{RW,IW} = \beta_{RB,IB} - \beta_{RW,IB}$ (i.e., that the difference in assessments of blacks and whites by black interviewers is equal to the difference in assessments of blacks and whites by white interviewers) indicates the presence of bias, it does not indicate whether the bias lies in black interviewers' assessments of white respondents, in white interviewers' assessments of black respondents, or in some combination of both.

A final set of models includes interactions between respondent and

interviewer race, as well as interviewer fixed effects. The inclusion of interviewer fixed effects accounts for unobserved interviewer characteristics that may be associated with both interviewer race and the assessed values of the outcome variables. This better isolates the effects of the respondent-interviewer race interactions and decreases the likelihood that such results stem from interviewer characteristics other than race. The model is of the form

$$y_{RI} = \beta_i + \beta_{RB,IW}I(R \text{ Black}, I \text{ White}) \\ + \beta_{RB,IB}I(R \text{ Black}, I \text{ Black}) + \mathbf{X}\alpha + \varepsilon, \quad (4)$$

where the fixed effects are represented by the intercept β_i , which varies across interviewers. The inclusion of interviewer fixed effects necessitates excluding the city dummies (which do not vary for individual interviewers) and one of the racial interaction terms. The parameter $\beta_{RB,IW}$ measures the difference in outcomes across black and white respondents who are assessed by white interviewers. The parameter $\beta_{RB,IB}$ measures the difference in outcomes across black and white respondents who are assessed by black interviewers. The null hypothesis of no racial bias is simply that $\beta_{RB,IW} = \beta_{RB,IB}$.

Results

Are racial differences in outcomes explained by sociodemographic factors and maternal characteristics?—The descriptive statistics presented in table 1 suggest that, without accounting for any other factors, blacks and whites differ substantially in terms of both the self-reported and interviewer-observed parenting measures. However, these racial differences may be accounted for by other household and parental characteristics that vary across racial groups. Estimates of equation (1) in table 4 (parenting measures) and table 5 (measures of maternal and child characteristics) examine whether this is the case.

Table 4 shows results for the mother-assessed and interviewer-assessed parenting measures. Results in the top panel of table 4 (model 1) indicate that, controlling only for city of residence, black mothers are more likely than white mothers to score in both the seventy-fifth and ninetieth percentile on all of these measures except psychological aggression. However, once sociodemographic characteristics are added (model 2), many of these effects become much smaller or vanish entirely. Black mothers are still statistically significantly more likely to score in the seventy-fifth percentile for lack of nonviolent discipline and lack of warmth, as well as in the ninetieth percentile for physical assault. Yet, the race effect for all other measures becomes small and statistically insignificant. Model 3 adds measures of maternal depression and risky prenatal behaviors. It is interesting to note that these measures, the

Table 4
MATERNAL RACE AND PARENTING

CUT AT PERCENTILE	ASSESSED BY MOTHER						ASSESSED BY INTERVIEWER			
	Lack of Non-violent Discipline		Psychological Aggression		Physical Assault		Harshness		Lack of Warmth	
	75th	90th	75th	90th	75th	90th	75th	90th	75th	90th
Model 1 (controls for mother's race and city):										
Mother is black	.139**	.052**	.041	.015	.062*	.041*	.078**	.068**	.158**	.064**
	(.028)	(.019)	(.029)	(.021)	(.029)	(.018)	(.024)	(.020)	(.029)	(.018)
Model 2 (add household sociodemographic characteristics):										
Mother is black	.093**	.019	.025	-.016	.055	.040*	.021	.023	.079*	.024
	(.031)	(.021)	(.032)	(.023)	(.032)	(.020)	(.026)	(.022)	(.031)	(.020)
<i>F</i> -test (sociodemographics) ^a	.009	.000	.047	.001	.066	.241	.000	.000	.000	.000
Model 3 (add maternal depression and risky behaviors):										
Mother is black	.085**	.013	.029	-.009	.051	.044*	.023	.022	.079*	.023
	(.032)	(.021)	(.032)	(.023)	(.033)	(.021)	(.027)	(.022)	(.032)	(.020)
<i>F</i> -test (sociodemographics) ^a	.002	.000	.110	.042	.138	.412	.000	.000	.000	.002
<i>F</i> -test (maternal depression and behaviors) ^a	.055	.056	.000	.000	.041	.269	.798	.928	.485	.758

NOTE.—1,417 observations; ordinary least squares regressions. The dependent variable equals one if above the relevant percentile cut point for each parenting measure, otherwise zero. Coefficients and standard errors (in parentheses) shown in rows marked "Mother is black"; *p* values from *F*-tests are for the hypothesis that the variables are jointly insignificant. The controls for sociodemographic characteristics, maternal depression, and maternal behaviors are those shown in table 2.

^a Denotes *p* value.

* *p* < .05.

** *p* < .01.

Table 5

MATERNAL RACE AND MATERNAL AND CHILD BEHAVIORS AND CHARACTERISTICS ASSESSED BY INTERVIEWER

CUT AT PERCENTILE	LACK OF MATERNAL VERBAL AND SOCIAL SKILLS		LACK OF MATERNAL UNDERSTANDING AND ATTENTION		MATERNAL HOSTILITY AND SUSPICION	PROBLEMS WITH CHILD'S APPEARANCE		PROBLEMS WITH CHILD'S BEHAVIOR	
	75th	90th	75th	90th	90th	75th	90th	75th	90th
Model 1 (controls for mother's race and city): Mother is black	.080** (.023)	.035* (.016)	.145** (.029)	.048** (.017)	.030 (.016)	.059* (.029)	.017 (.016)	.007 (.029)	.011 (.019)
Model 2 (add household sociodemographic characteristics): Mother is black	.036 (.025)	.011 (.018)	.044 (.031)	.007 (.019)	.019 (.017)	-.043 (.031)	-.024 (.018)	-.050 (.032)	-.022 (.021)
<i>F</i> -test (sociodemographics) ^a	.000	.059	.000	.000	.000	.000	.000	.000	.005
Model 3 (add maternal depression and risky behaviors): Mother is black	.039 (.025)	.017 (.018)	.047 (.031)	.010 (.019)	.024 (.017)	-.044 (.031)	-.018 (.018)	-.045 (.032)	-.022 (.021)
<i>F</i> -test (sociodemographics) ^a	.001	.177	.000	.005	.000	.000	.000	.000	.007
<i>F</i> -test (maternal depression and behaviors) ^a	.169	.047	.327	.208	.311	.989	.188	.754	.984

NOTE.—1,417 observations; ordinary least squares regressions. The dependent variable equals one if above the relevant percentile cut point for each measure, otherwise zero. Coefficients and standard errors (in parentheses) are shown in rows marked "Mother is black"; *p* values from *F*-tests are for the hypothesis that the variables are jointly insignificant. The controls for sociodemographic characteristics, maternal depression, and maternal behaviors are those shown in table 2.

^a Denotes *p* value.

* *p* < .05.

** *p* < .01.

values of which are unknown to interviewers and might be difficult for them to assess during the course of the interview, are often associated with mothers' reports of their own parenting behaviors but are unrelated to interviewers' assessments. However, the addition of these measures has little additional effect on the estimated race coefficients, even for those models in which these characteristics are jointly statistically significant. The results in table 5, for the interviewer-assessed maternal and child behaviors and characteristics, are similar to those in table 4. All initially significant racial differences in the outcomes disappear once sociodemographic variables are included in the models.

On the whole, these results indicate that blacks are more likely than whites to have extreme scores on many parenting measures, whether mother-reported or interviewer-assessed. The same is true of some maternal and child behaviors and characteristics measures, regardless of who makes the assessment. However, a considerable proportion of the racial difference on these measures can be explained by observable sociodemographic characteristics, while little of the difference can be explained by maternal depression and risky behaviors.¹⁸ Furthermore, the results in tables 4 and 5 suggest that sociodemographic characteristics do a better job of explaining black-white differences in the maternal and child behaviors and characteristics than in the parenting measures. Sociodemographic characteristics also do a better job of explaining racial differences in the interviewer-assessed measures than in the mother-reported measures.

Does the race of the interviewer affect assessments?—Estimates of equation (2), which add an indicator for the race of the interviewer, are presented in the top panels of table 6 (for parenting measures) and table 7 (for maternal and child behaviors and characteristics). Overall, these results indicate that the race of the interviewer matters for many more of the interviewer-assessed items than the items based on maternal reports. In table 6, the race of the interviewer is associated with only one of the self-reported discipline measures (lack of nonviolent discipline at the ninetyeth percentile). In contrast, it is associated with all of the interviewer-assessed measures. Black interviewers are significantly less likely to rate respondents as falling into the seventy-fifth and ninetyeth percentiles for harshness or lack of warmth. Similar results are obtained in table 7; for nine of the 10 measures presented, black interviewers rate parents or children with significantly lower (i.e., less-problematic) scores. The effects of the interviewer's race are generally large and are often greater in absolute value than the coefficients for the mother's own race.

Although this finding does not indicate the presence of racial bias, it does have implications for racial differences in reports of child maltreatment; if families are more likely to come into contact with reporters of the same race, then black families may be less harshly judged, on average, than white families. Furthermore, the results have implications

Table 6

TESTS FOR INTERVIEWER EFFECTS: PARENTING MEASURES

CUT AT PERCENTILE	ASSESSED BY MOTHER						ASSESSED BY INTERVIEWER			
	Lack of Non-violent Discipline		Psychological Aggression		Physical Assault		Harshness		Lack of Warmth	
	75th	90th	75th	90th	75th	90th	75th	90th	75th	90th
Equation 2 (controls for mother and interviewer race):										
Mother is black	.085**	.012	.029	-.009	.051	.044*	.024	.022	.084**	.026
	(.032)	(.021)	(.032)	(.023)	(.033)	(.021)	(.027)	(.022)	(.030)	(.019)
Interviewer is black	.023	.068**	.020	.030	-.011	.012	-.101**	-.056*	-.380**	-.189**
	(.032)	(.022)	(.033)	(.023)	(.033)	(.021)	(.027)	(.022)	(.031)	(.020)
Equation 3 (interactions among mother and interviewer race):										
Mother is black and interviewer is white	.084*	.000	.009	-.010	.035	.054*	.057	.058*	.106**	.062**
	(.039)	(.026)	(.040)	(.028)	(.040)	(.025)	(.033)	(.027)	(.037)	(.024)
Mother is black and interviewer is black	.107*	.073*	.037	.020	.031	.062*	-.058	-.013	-.284**	-.142*
	(.047)	(.031)	(.048)	(.034)	(.048)	(.030)	(.039)	(.032)	(.044)	(.029)
Mother is white and interviewer is black	.021	.044	-.020	.029	-.044	.034	-.035	.017	-.334**	-.115**
	(.057)	(.038)	(.058)	(.041)	(.059)	(.037)	(.048)	(.039)	(.054)	(.035)
Test for bias ^a	.966	.439	.405	.974	.500	.470	.096	.023	.296	.010
Equation 4 (interactions among mother and interviewer race, interviewer fixed effects):										
Mother is black and interviewer is white	.088*	.007	.010	-.006	.037	.057*	.045	.055*	.077*	.053*
	(.040)	(.026)	(.040)	(.029)	(.041)	(.026)	(.032)	(.027)	(.035)	(.024)
Mother is black and interviewer is black	.081	.028	.053	-.004	.084	.022	-.037	-.039	.065	-.012
	(.048)	(.032)	(.049)	(.035)	(.050)	(.032)	(.039)	(.033)	(.043)	(.029)
Test for bias ^a	.905	.590	.478	.961	.440	.365	.090	.019	.826	.068

NOTE.—Coefficients and standard errors (in parentheses) from ordinary least squares regressions are shown. All models include the sociodemographic and maternal characteristics listed in table 2. Equations 2 and 3 include city dummies. Equation 4 includes interviewer fixed effects.

^a Denotes *p* value.

* *p* < .05.

** *p* < .01.

Table 7

TESTS FOR INTERVIEWER EFFECTS: MATERNAL AND CHILD BEHAVIORS AND CHARACTERISTICS ASSESSED BY INTERVIEWER

CUT AT PERCENTILE	LACK OF MATERNAL VERBAL AND SOCIAL SKILLS		LACK OF MATERNAL UNDERSTANDING AND ATTENTION		MATERNAL HOSTILITY AND SUSPICION	PROBLEMS WITH CHILD'S APPEARANCE		PROBLEMS WITH CHILD'S BEHAVIOR	
	75th	90th	75th	90th	90th	75th	90th	75th	90th
Equation 2 (controls for mother and interviewer race):									
Mother is black	.042 (.024)	.018 (.018)	.048 (.031)	.012 (.019)	.026 (.017)	-.042 (.031)	-.017 (.018)	-.043 (.032)	-.021 (.021)
Interviewer is black	-.234** (.025)	-.124** (.018)	-.058 (.032)	-.098** (.019)	-.096** (.018)	-.132** (.032)	-.089** (.018)	-.105** (.033)	-.060** (.022)
Equation 3 (interactions among mother and interviewer race):									
Mother is black and interviewer is white	.089** (.030)	.046* (.022)	.060 (.039)	.025 (.023)	.043* (.021)	-.085* (.039)	-.021 (.022)	-.016 (.040)	.007 (.026)
Mother is black and interviewer is black	-.165** (.036)	-.090** (.026)	-.004 (.046)	-.079** (.028)	-.060* (.026)	-.199** (.046)	-.109** (.027)	-.133** (.048)	-.065* (.031)
Mother is white and interviewer is black	-.138** (.044)	-.068* (.032)	-.034 (.056)	-.072* (.034)	-.060 (.031)	-.219** (.056)	-.098** (.032)	-.050 (.058)	-.004 (.038)
Test for bias ^a	.007	.034	.596	.349	.163	.060	.753	.241	.071
Equation 4 (interactions among mother and interviewer race, interviewer fixed effects):									
Mother is black and interviewer is white	.067* (.030)	.035 (.022)	.045 (.038)	.009 (.023)	.031 (.021)	-.086* (.038)	-.020 (.022)	-.031 (.041)	-.005 (.027)
Mother is black and interviewer is black	.013 (.036)	-.016 (.026)	.032 (.046)	-.006 (.028)	.019 (.026)	.041 (.046)	-.012 (.027)	-.080 (.049)	-.076* (.032)
Test for bias ^a	.217	.112	.815	.656	.716	.025	.824	.422	.074

NOTE.—Coefficients and standard errors (in parentheses) from ordinary least squares regressions are shown. All models include the sociodemographic and maternal characteristics listed in table 2. Equations 2 and 3 include city dummies. Equation 4 includes interviewer fixed effects.

^a Denotes p value.

* $p < .05$.

** $p < .01$.

for survey implementation: estimates of racial differences in outcomes may be quite different depending on whether or not interviewers are matched to respondents on the basis of their races.

Does the race of the interviewer relative to the race of the mother affect assessments?—The lower panels of tables 6 and 7 present estimates of equations (3) and (4). These equations include interactions between the mother's and interviewer's races. These results provide the basis for the tests of racial bias. The results in table 6 indicate that there is no evidence of racial bias in the maternal-assessed measures of parenting. Consistent with the results in the top panel, black mothers are somewhat more likely to report using few nonviolent discipline strategies and are more likely to fall into the ninetieth percentile for physical assault. However, this is true regardless of whether they are assigned to black or white interviewers. Similarly, the reports of white mothers do not depend on the race of the interviewer to whom those mothers are assigned. For each of the three maternal-assessed measures (lack of nonviolent discipline, psychological aggression, and physical assault), the hypothesis of no racial bias cannot be rejected. These findings indicate that racial interactions do not influence how mothers respond to questions about their parenting strategies.

Results for some of the interviewer-assessed parenting items are quite different. Consider, for example, the results for harshness (table 6). Blacks who were interviewed by whites have higher (worse) scores than blacks who had black interviewers. Differencing the coefficients for mother is black and interviewer is white, and mother is black and interviewer is black, indicates that black mothers are 11.5 percentage points more likely to score above the seventy-fifth percentile and 7.1 percentage points more likely to score above the ninetieth percentile, if they are interviewed by whites rather than by blacks. In part, this could result from the earlier finding that black interviewers are more lenient when assessing all mothers, black and white. However, the results for white mothers indicate that this is not the case. White mothers interviewed by whites are only 3.5 percentage points more likely to fall above the seventy-fifth percentile and are 1.7 percentage points less likely to fall above the ninetieth percentile than are white mothers interviewed by blacks. Both of these differences are statistically insignificant. The hypothesis that no racial bias exists in the interviewer assessment for harshness is rejected at the 9.6 percent level when using the seventy-fifth percentile cut point, and at the 2.3 percent level when using the ninetieth percentile cut point.¹⁹ Similar conclusions for harshness can be drawn from the estimates of equation (4). Those estimates include interviewer fixed effects.

Although the measure of maternal harshness shows clear evidence of racial bias, this is less true for lack of maternal warmth. The hypothesis of no racial bias is rejected only when the ninetieth percentile threshold is used. In this case, black mothers are 20.4 percentage points more

likely to score above the ninetieth percentile when interviewed by whites than when by blacks, and white mothers are 11.5 percentage points more likely to score above the ninetieth percentile when interviewed by whites than when by blacks. The difference-in-difference (8.9 percentage points) is statistically significant at the 1 percent level. Again, the fixed effects estimates yield similar results.

There is some evidence of racial bias in interviewers' assessments of parenting, but there is less clear-cut evidence of bias in maternal and child behaviors and characteristics. When interviewer fixed effects are not included, there is evidence of racial bias in the assessments of maternal verbal and social skills (at both the seventy-fifth and ninetieth percentiles), of problems with the child's appearance (at the seventy-fifth percentile), and of problems with the child's behavior (at the ninetieth percentile). For lack of maternal verbal and social skills and problems with child behavior, the pattern of coefficients is the same as that for the parenting measures described above; black mothers are more likely to receive high scores when interviewed by whites than when interviewed by blacks, and this gap is larger than is observed for white mothers. For problems with the child's appearance at the seventy-fifth percentile, a somewhat different pattern is observed; white interviewers are more likely to note problems with the appearance of white children than black children, whereas black interviewers are slightly more likely to note problems with the appearance of black children than with white children. However, when interviewer fixed effects are included, the standard errors become sufficiently large such that the hypothesis of no racial bias can be rejected only for problems with the child's appearance (seventy-fifth percentile cut point) and child's behavior (at the ninetieth percentile cut point). There is no evidence of racial bias in assessments of lack of maternal understanding and attention or maternal hostility and suspicion.

The finding of racial bias in some measures of the child's behavior may provide a clue as to the source of racial bias in the assessments of harshness discussed above. At least according to interviewers' (possibly biased) assessments, black children display fewer behavior problems during interviews with black interviewers than with their white counterparts. The behavior of white children is rated the same regardless of the race of the interviewer. If these differences in child behavior are in fact genuine, then black parents may have had reason to display more punitive behavior during interviews conducted by whites. Under this interpretation, the findings of racial bias could be due to changes in the child's behavior that result from the racial composition of the interview. There is an alternative explanation, however, and it is difficult to distinguish from the first; black and white interviewers may in fact judge parents and children with similar behaviors in different ways. Under either interpretation, it is still the case that judgments about

several dimensions of parenting and child behavior are influenced by the race of the interviewer relative to that of the family being assessed.

A final set of results, reported in table 8, examines racial bias in the individual items that make up the scales for harshness and lack of warmth. As discussed above, theories of racial bias in the formation of judgments (in a context of imperfect information about true behaviors) imply that racial bias (if it exists) should be more common in situations in which there is more ambiguity and potential for judgment calls by interviewers than in more clear-cut situations. The items in the scales for harshness and lack of warmth vary in the degree to which judgment is required. For example, one item included in the harshness scale, whether the parent did not slap or spank the child during the visit, leaves little room for an interviewer's judgment. However, another item, whether the parent expressed annoyance with or hostility toward the child during the visit, requires the interviewer to make a more subjective assessment. Table 8 presents results from regressions of the form of equation (3). These models include complete sets of interactions between the interviewer's and mother's race but use the individual items in the harshness and lack of warmth scales as dependent variables. Each item is coded as a dichotomous variable, with a value of one indicating a better (less harsh or more warm) outcome. These estimates provide evidence of whether items that require more judgment on the part of interviewers more frequently show evidence of racial bias.

The results in table 8 provide support for the hypothesis that racial bias increases with the amount of judgment required by interviewers. This is most clear for the items that make up the harshness scale. The statistical tests indicate that there is racial bias for the items assessing parental annoyance with or hostility toward a child and parental scolding or criticism of the child. In contrast, the tests do not find bias for more objective items, such as whether the parent did not shout at, slap, spank, interfere with, or restrict the child. That said, it should be noted that the general pattern is consistent with racial bias for all items: black mothers interviewed by whites are less likely to receive positive ratings than white mothers interviewed by whites; black mothers interviewed by blacks are rated to be very similar to white mothers interviewed by blacks. Recall that the results presented in table 6 show evidence of bias on the harshness scale at both the seventy-fifth and ninetieth percentiles.

Patterns are similar for lack of warmth. However, it is more difficult to break these items into those that require more or less judgment. The hypothesis that there is no racial bias in the assessment is rejected for one item that clearly seems to require interviewer judgment: whether the parent's voice conveys positive feelings toward the child. It is not rejected for the more objective item reporting whether the parent caressed or kissed the child at least once during the interview. Overall,

Table 8

TESTS FOR INTERVIEWER EFFECTS: SUBITEMS IN THE HARSHNESS AND LACK OF WARMTH SCALES

	Mother Is Black and Interviewer Is White	Mother Is Black and Interviewer Is Black	Mother Is White and Interviewer Is Black	Test for Bias ^a
Items included in the harshness scale:				
Parent did not shout at child	-.053 (.031)	.018 (.037)	.012 (.045)	.190
Parent did not express annoyance or hostility toward child	-.088** (.033)	.124*** (.039)	.104* (.048)	.022
Parent neither slapped nor spanked child	-.025 (.017)	-.016 (.020)	.004 (.024)	.843
Parent did not scold or criticize child	-.068* (.032)	-.011 (.038)	-.052 (.046)	.012
Parent did not interfere or restrict child more than three times	-.042 (.035)	.003 (.043)	.006 (.052)	.456
Items included in the lack of warmth scale:				
Parent spontaneously vocalized to child twice	-.034 (.027)	.111*** (.032)	.108** (.039)	.335
Parent responded verbally to child's vocalizations	-.060* (.024)	.101*** (.028)	.086** (.034)	.028
Parent told child the name of an object or person during visit	-.109** (.040)	.214*** (.049)	.277*** (.059)	.435
Parent spontaneously praised child at least twice	-.131*** (.035)	.233*** (.042)	.260*** (.050)	.039
Parent's voice conveys positive feelings toward child	-.067** (.025)	.095*** (.030)	.078* (.036)	.020
Parent caressed or kissed child at least once	-.113** (.039)	.178*** (.046)	.230*** (.056)	.276

NOTE.—Each row shows coefficients and standard errors (in parentheses) from a single ordinary least squares regression, using the sample in table 2. All dependent variables are coded one if the interviewer judged the statement to be true, otherwise zero. All models include the sociodemographic and maternal characteristics listed in table 2 and a set of city dummies.

^a Denotes p value.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

the hypothesis that there is no racial bias in the assessment is rejected for three of the six items.

Conclusions

The families in this study were not the subjects of CPS investigations, nor did the interviewers have the job of making assessments of maltreatment. However, the results provide several insights into possible sources of reporting bias within the child welfare system. The empirical analyses focus on three major areas of inquiry: (1) whether there are racial differences in mother-reported and interviewer-assessed measures of parenting; (2) whether differences in parenting across racial groups are accounted for by household and maternal characteristics; and (3) whether the race of the interviewer, alone and relative to the race of the mother, influences interviewer assessments of parenting.

The first major finding of these analyses is that there are racial differences in many measures of parenting. The presence of these differences among both self-reported and interviewer-assessed items indicates that they are not solely the product of bias on the part of interviewers. However, a large portion of these racial differences are explained by measures of socioeconomic status that are correlated with race. These results have implications for the possibility that reporting bias may be driven by statistical discrimination. As discussed above, the race of families is likely to be more easily observed by potential reporters than family income, maternal education, or the other measures included in this study's set of sociodemographic variables. In a context of incomplete information, race may be used as a proxy for low socioeconomic status in making judgments about parenting, and, absent information on non-racial characteristics, potential reporters may attribute poor behavior to all blacks, regardless of their socioeconomic status.

Second, there is evidence that race matters when individuals assess the parenting behaviors of others. There are systematic differences in how black and white interviewers assess parenting and parent and child behaviors and characteristics. For the measures of harshness, lack of warmth, problems with the child's appearance, and problems with child's behavior, there is evidence of racial bias in assessments. That an interviewer's observations of parent-child interactions are affected by the interviewer's race calls into question the idea that reports of child maltreatment are color-blind.

Although the results of this study indicate that race influences judgments, they do not provide information on the extent to which higher rates of maltreatment reports among black children result from racial bias. The finding of racial differences in maternal self-reported parenting behaviors suggests that higher report rates for blacks relative to whites may, in part, reflect genuine differences in parenting behaviors. Such

differences may be driven by characteristics that are correlated with race. However, even this conclusion must be treated with caution because none of the parenting measures is extreme enough to warrant CPS reports, and it is possible that there are no racial differences in measures of actual maltreatment.²⁰ Similarly, although there is evidence of racial bias in judgments about parenting, it is not clear whether similar biases would appear for judgments about more extreme parenting behaviors or if racial bias in judgments would translate into racially biased reporting behaviors. More research is required to answer these questions.

These results also have implications for researchers collecting data on parenting. The findings underscore the importance of considering interviewer race, as well as the relationship between the interviewer's and the respondent's races, when studying the determinants of parenting. We are not advocating for interviewer-respondent matching based on race: on the contrary, the finding that black and white interviewers assess families and children differently implies that matching could yield very misleading results. However, it is important to examine whether conclusions are sensitive to the assignment of interviewers to respondents.

Appendix

Construction of Outcome Measures

The items listed below were drawn from the Fragile Families and Child Wellbeing Study's In-Home Longitudinal Survey of Pre-School Age Children questionnaire. The number and letter preceding each item represent the section (letter) and item (number) for each variable. Many of the items listed here are paraphrased, rather than directly quoted from the questionnaire. In all cases, scales are coded such that higher scores are worse. While citations to the original sources of these items are included below, many of the items were adapted from their original form and may not perfectly match those found in the original sources.

Mother-Assessed Parenting Measures

These measures are drawn from subscales of the Parent-Child Conflict Tactics Scales (Straus et al. 1998). For each measure, individuals were assigned a score of 0 if they responded "this has never happened" or "yes, but not in the past year" for a particular item. They were assigned a score of 1 if they reported that the event occurred once in the year prior to the interview, 2 for twice, 4 for 3–5 times, 8 for 6–10 times, 15 for 11–20 times, and 25 for more than 20 times in the year prior to the interview. Scores on each item were then summed to compute total yearly frequency scores for each subscale.

Lack of Nonviolent Discipline (0–100 Points; Reverse Coded):

- J1. Explained why something was wrong
- J5. Gave him or her something else to do instead of what he or she was doing
- J12. Took away privileges from him or her
- J2. Put in time-out (or sent to room)

Psychological Aggression (0–125 Points):

- J6. Shouted, yelled, or screamed at
- J10. Threatened to spank or hit but did not actually do it
- J8. Swore or cursed at
- J14. Called him or her dumb or lazy or some other name like that
- J9. Said would send him or her away or would kick him or her out of the house

Physical Assault (0–125 Points):

- J7. Spanked him or her on the bottom with bare hand
- J4. Hit him or her on the bottom with something like a belt, hairbrush, a stick, or some other hard object
- J11. Slapped him or her on the hand, arm, or leg
- J13. Pinched him or her
- J3. Shook him or her

Interviewer-Assessed Parenting Measures

These measures are drawn from subscales of the HOME (Caldwell and Bradley 1984). For each, one point was assigned for each affirmative response. The number of affirmative responses was then summed to create a total score for each subscale.

Harshness (0–5 Points; Reverse Coded):

- T10. Parent did not shout at child (e.g., did not raise voice above level required by distance between mother and child)
- T11. Parent did not express annoyance with or hostility toward child
- T12. Parent neither slapped nor spanked child during the visit
- T13. Parent did not scold or criticize child during visit
- T14. Parent did not interfere or restrict child more than three times (does not include protecting child from harm)

680 Social Service Review

Lack of Warmth (0–6 Points; Reverse Coded):

- T1. Parent spontaneously vocalized to child twice
- T2. Parent responded verbally to child’s vocalizations
- T3. Parent told child the name of an object or person during visit
- T7. Parent spontaneously praised child at least twice
- T8. Parent’s voice conveys positive feelings toward child
- T9. Parent caressed or kissed child at least once

Interviewer-Assessed Maternal and Child Behaviors and Characteristics

Lack of Maternal Verbal and Social Skills (0–3 Points; Reverse Coded):

This measure is based on HOME items (Caldwell and Bradley 1984). One point was assigned for each affirmative response. The number of affirmative responses was then summed to create a total score for the subscale.

- T4. Parent’s speech was distinct and audible
- T5. Parent initiated verbal exchanges with visitor
- T6. Parent conversed freely and easily

Lack of Maternal Attention and Understanding (0–12 Points; Reverse Coded):

This measure is based on the interviewer’s assessment of the mother over the course of the entire interview. Each interviewer completed these items immediately following the interview. Each item was scored on a zero-to-three-point scale ranging from “poor” to “excellent” (for item V5, the scale ranged from “very uncooperative” to “very cooperative”). Scores on each item were then summed to create a total score for the subscale.

- V2. Respondent’s attention to interviewer was . . .
- V3. Respondent’s understanding of the questions was . . .
- V4. Respondent’s ability to articulate answers was . . .
- V5. Respondent’s cooperation throughout most of the interview was . . .

Maternal Hostility and Suspicion (0–10 Points):

This measure is based on the interviewer’s assessment of the mother over the course of the entire interview. Each interviewer completed these items immediately following the interview. Each item was scored on a zero-to-two-point scale ranging from “no” to “very.” Scores on each item were then summed to create a total score for the subscale.

- V6A. Did respondent appear suspicious?
- V6B. Did respondent appear uncommunicative?
- V6C. Did respondent appear anxious or nervous?
- V6D. Did respondent appear hostile?
- V6E. Did respondent appear to be on drugs?

Problems with Child's Appearance (0–22 Points; Reverse Coded Where Appropriate):

One point was assigned for each affirmative response for 12 items (S1_1 through S1_11 and S5). Item S2 is coded on a four-point scale; S3 and S4 are coded on three-point scales. Scores on each item were summed to create a total score for the subscale.

- S1_1. Child's clothing is dirty or unkempt
- S1_2. Child's clothing dirty due to playing or eating
- S1_3. Child is in soiled diaper
- S1_4. Child may be in soiled diaper
- S1_5. Clothing is worn, but mended or not ripped or torn
- S1_6. Clothing is worn, but not mended, obvious rips or tears
- S1_7. Clothing is too tight for comfortable fit
- S1_8. Clothing is too large
- S1_9. Clothing is too lightweight for indoor temperature (underdressed)
- S1_10. Clothing is too warm for indoor temperature (overdressed)
- S1_11. Other negative conditions not specified
- S2. Child washed or recently bathed
- S3. Child's hair combed and clean
- S4. Child emits no body or mouth odor
- S5. Anything else about child's clothing or hygiene that is problematic

Problems with Child's Behavior (0–20 Points; Reverse Coded Where Appropriate):

This measure is based on the interviewer's assessment of the child over the course of the entire interview. Each interviewer completed these items immediately following the interview. Each item was scored on a zero-to-four-point scale. Scores on each item were then summed to create a total score for the subscale.

- U1. Did the child display positive emotions during the visit?
- U2. Did the child display negative emotions during the visit?
- U3. How persistent was the child when completing the Peabody Picture Vocabulary Test—Revised (PPVT) or Test de Vocabulario en Imagenes Peabody (TVIP)?
- U4. How cooperative was the child when completing the PPVT or TVIP?
- U5. How cooperative was the child while being weighed and measured?

Lack of maternal attention and understanding:												
Black	29.35	6.39	8.06	28.61	9.72	3.98	5.74	4.72	1.76	.74	.93	79.14 (.000)
White	52.82	8.31	4.75	22.26	3.56	2.67	3.26	1.78	.59	.00	.00	
Maternal hostility and suspicion:												
Black	86.48	7.13	2.87	1.57	1.48	.19	.28					
White	95.21	1.80	1.50	.60	.30	.30	.30					
Problems with child's appearance:												
Black	74.44	8.70	4.72	5.19	2.41	1.48	.83	1.20	.74	.00	.28	24.01 (.013)
White	84.27	7.12	3.26	2.08	1.19	.89	.89	.00	.00	.30	.00	
Problems with child's behavior:												
Black	14.26	8.70	12.78	10.19	8.06	7.96	5.83	4.54	3.24	3.61	18.61	24.40 (.225)
White	16.91	9.20	16.91	11.57	9.20	7.72	2.97	5.04	3.26	4.75	12.46	

NOTE.—For each measure, each row shows the frequency distribution for the sample of black and white mothers. The last column shows results from a χ^2 test of the null hypothesis that the distributions are the same for black and white mothers (*p* values are in parentheses).

References

- Aigner, Dennis J., and Glen G. Cain. 1977. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review* 30 (2): 175–87.
- Ards, Sheila D., Samuel L. Myers Jr., Chanjin Chung, Allan Malkis, and Brian Hagerty. 2001. "Racial Reporting Bias and Child Maltreatment." Paper presented at the International Society for the Prevention of Child Abuse and Neglect, VIII European Conference on Child Abuse and Neglect, Istanbul, Turkey, August 24–27.
- . 2003. "Decomposing Black-White Differences in Child Maltreatment." *Child Maltreatment* 8 (2): 112–21.
- Balsa, Ana I., and Thomas G. McGuire. 2001. "Statistical Discrimination in Health Care." *Journal of Health Economics* 20 (6): 881–907.
- . 2003. "Prejudice, Clinical Uncertainty and Stereotyping as Sources of Health Disparities." *Journal of Health Economics* 22 (1): 89–116.
- Berger, Lawrence M., and Jeanne Brooks-Gunn. 2005. "Socioeconomic Status, Parenting Knowledge and Behaviors, and Perceived Maltreatment of Young Low-Birth-Weight Children." *Social Service Review* 79 (2): 237–67.
- Berlin, Lisa J., Jeanne Brooks-Gunn, Donna Spiker, and Martha J. Zaslow. 1995. "Examining Observational Measures of Emotional Support and Cognitive Stimulation in Black and White Mothers of Preschoolers." *Journal of Family Issues* 16 (5): 664–86.
- Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94 (4): 991–1013.
- Bradley, Robert H. 1993. "Children's Home Environments, Health, Behavior, and Intervention Efforts: A Review Using the HOME Inventory as a Marker Measure." *Genetic, Social, and General Psychology Monographs* 119 (4): 437–90.
- Bradley, Robert H., Robert F. Corwyn, Harriette Pipes McAdoo, and Cynthia Garcia Coll. 2001. "The Home Environments of Children in the United States: Part 1; Variations by Age, Ethnicity, and Poverty Status." *Child Development* 72 (6): 1844–67.
- Caldwell, Bettye M., and Robert H. Bradley. 1984. *Administration Manual: HOME Observation for Measurement of the Home Environment*. Rev. ed. Little Rock: University of Arkansas.
- Chasnoff, Ira J., Harvey J. Landress, and Mark E. Barrett. 1990. "The Prevalence of Illicit-Drug or Alcohol Use during Pregnancy and Discrepancies in Mandatory Reporting in Pinellas County, Florida." *New England Journal of Medicine* 322 (17): 1202–6.
- Courtney, Mark E., Richard P. Barth, Jill Duerr Berrick, Devon Brooks, Barbara Needell, and Linda Park. 1996. "Race and Child Welfare Services: Past Research and Future Directions." *Child Welfare* 75 (2): 99–137.
- Dietz, Tracy L. 2000. "Disciplining Children: Characteristics Associated with the Use of Corporal Punishment." *Child Abuse and Neglect* 24 (12): 1529–42.
- Drake, Brett, and Susan Zuravin. 1998. "Bias in Child Maltreatment Reporting: Revisiting the Myth of Classlessness." *American Journal of Orthopsychiatry* 68 (2): 295–304.
- Egu, Chizoma Linda, and David J. Weiss. 2003. "The Role of Race and Severity of Abuse in Teachers' Recognition or Reporting of Child Abuse." *Journal of Child and Family Studies* 12 (4): 465–74.
- Fiske, Susan T. 1998. "Stereotyping, Prejudice, and Discrimination." 357–411 in *The Handbook of Social Psychology*, edited by Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey. Vol. 2, 4th ed. New York: McGraw-Hill.
- Fluke, John D., Ying-Ying T. Yuan, John Hedderson, and Patrick A. Curtis. 2003. "Disproportionate Representation of Race and Ethnicity in Child Maltreatment: Investigation and Victimization." *Children and Youth Services Review* 25 (5–6): 359–73.
- Frazer, Ricardo A., and Uco J. Wiersma. 2001. "Prejudice versus Discrimination in the Employment Interview: We May Hire Equally, but Our Memories Harbour Prejudice." *Human Relations* 54 (2): 173–91.
- Fryer, Roland G., Jr., and Matthew O. Jackson. 2003. "Categorical Cognition: A Psychological Model of Categories and Identification in Decision Making." Working Paper no. 9579. National Bureau of Economic Research, Cambridge, MA.
- Fulgini, Allison Sidle, Wen-Jui Han, and Jeanne Brooks-Gunn. 2004. "The Infant-Toddler HOME in the 2nd and 3rd Years of Life." *Parenting: Science and Practice* 4 (2–3): 139–60.
- Garland, Ann F., Elissa Ellis-MacLeod, John A. Landsverk, William Ganger, and Ivory

- Johnson. 1998. "Minority Populations in the Child Welfare System: The Visibility Hypothesis Reexamined." *American Journal of Orthopsychiatry* 68 (1): 142–46.
- Gil, David G. 1970. *Violence against Children: Physical Child Abuse in the United States*. Cambridge, MA: Harvard University Press.
- Goldberg, Pinelopi Koujianou. 1996. "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey." *Journal of Political Economy* 104 (3): 622–54.
- Hampton, Robert L., and Eli H. Newberger. 1985. "Child Abuse Incidence and Reporting by Hospitals: Significance of Severity, Class, and Race." *American Journal of Public Health* 75 (1): 56–60.
- Harris, Marian S., and Mark E. Courtney. 2003. "The Interaction of Race, Ethnicity, and Family Structure with Respect to the Timing of Family Reunification." *Children and Youth Services Review* 25 (5–6): 409–29.
- Hines, Alice M., Kathy Lemon, Paige Wyatt, and Joan Merdinger. 2004. "Factors Related to the Disproportionate Involvement of Children of Color in the Child Welfare System: A Review and Emerging Themes." *Children and Youth Services Review* 26 (6): 507–27.
- Kessler, Ronald C., Gavin Andrews, Daniel Mroczek, Bedirhan Ustun, and Hans-Ulrich Wittchen. 1998. "The World Health Organization Composite International Diagnostic Interview Short-Form (CIDI-SF)." *International Journal of Methods in Psychiatric Research* 7 (4): 171–85.
- Kim, Hyoshin, Nazli Baydar, and April Greek. 2002. "Testing Conditions Influence the Race Gap in Cognition and Achievement Estimated by Household Survey Data." *Applied Developmental Psychology* 23 (5): 567–82.
- Lane, Wendy G., David M. Rubin, Ragin Monteith, and Cindy W. Christian. 2002. "Racial Differences in the Evaluation of Pediatric Fractures for Physical Abuse." *Journal of the American Medical Association* 288 (13): 1603–9.
- Leventhal, Tama, Anne Martin, and Jeanne Brooks-Gunn. 2004. "The EC-HOME across Five National Data Sets in the 3rd to 5th Year of Life." *Parenting: Science and Practice* 4 (2–3): 161–88.
- Levine, Murray, Howard J. Doueck, Jennifer B. Freeman, and Cheryl Compaan. 1996. "African American Families and Child Protection." *Children and Youth Services Review* 18 (8): 693–711.
- Lindsey, Duncan. 1994. *The Welfare of Children*. New York: Oxford University Press.
- Needell, Barbara, M. Alan Brookhart, and Seon Lee. 2003. "Black Children and Foster Care Placement in California." *Children and Youth Services Review* 25 (5–6): 393–408.
- Pakieser, Ruth A., Debra K. Starr, and Denise LeBaugh. 1998. "Nebraska School Nurses Identify Emotional Maltreatment of School-Age Children: A Replication of an Ohio Study." *Journal of the Society of Pediatric Nurses* 3 (4): 137–45.
- Phelps, Edmund. 1972. "The Statistical Theory of Racism and Sexism." *American Economic Review* 62 (4): 659–61.
- Prewett-Livingston, Amelia J., Hubert S. Feild, John G. Veres III, and Philip M. Lewis. 1996. "Effects of Race on Interview Ratings in a Situational Panel Interview." *Journal of Applied Psychology* 81 (2): 178–86.
- Reichman, Nancy E., Julien O. Teitler, Irwin Garfinkel, and Sara S. McLanahan. 2001. "Fragile Families: Sample and Design." *Children and Youth Services Review* 23 (4–5): 303–26.
- Roberts, Dorothy E. 2002. *Shattered Bonds: The Color of Child Welfare*. New York: Basic Civitas.
- Sagatun-Edwards, Inger, and Coleen Saylor. 2000. "Drug Exposed Infant Cases in Juvenile Court: Risk Factors and Court Outcomes." *Child Abuse and Neglect* 24 (7): 925–37.
- Straus, Murray A., and Richard J. Gelles, eds. 1990. *Physical Violence in American Families: Risk Factors and Adaptations to Violence in 8,145 Families*. New Brunswick, NJ: Transaction.
- Straus, Murray A., Sherry L. Hamby, David Finkelhor, David W. Moore, and Desmond Runyan. 1998. "Identification of Child Maltreatment with the Parent-Child Conflict Tactics Scales: Development and Psychometric Data for a National Sample of American Parents." *Child Abuse and Neglect* 22 (4): 249–70.
- Straus, Murray A., and Anita K. Mathur. 1996. "Social Change and the Trends in Approval of Corporal Punishment by Parents from 1968 to 1994." 91–105 in *Family Violence against Children: A Challenge for Society*, edited by Detlev Frehsee, Wiebke Horn, and Kai-D Bussmann. New York: de Gruyter.
- Sugland, Barbara W., Martha J. Zaslow, Judith R. Smith, Jeanne Brooks-Gunn, Deborah

- Coates, Connie Blumenthal, Kristin A. Moore, Terri Griffin, and Robert H. Bradley. 1995. "The Early Childhood HOME Inventory and HOME-Short Form in Differing Racial/Ethnic Groups: Are There Differences in Underlying Structure, Internal Consistency of Subscales, and Patterns of Protection?" *Journal of Family Issues* 16 (5): 632-63.
- U.S. Bureau of the Census. 2001. *Statistical Abstract of the United States, 2000*. Washington, DC: U.S. Bureau of the Census.
- . 2004. "Age and Sex of All People, Family Members and Unrelated Individuals Iterated by Income-to-Poverty Ratio and Race." Table pov01 in "Current Population Survey, 2004: Annual Social and Economic Supplement," revised November 22. Washington, DC: U.S. Bureau of the Census. Available at <http://pubdb3.census.gov/macro/032004/pov/toc.htm>.
- U.S. Department of Health and Human Services, Administration on Children, Youth, and Families. 2002. *Child Maltreatment 2000*. Washington, DC: U.S. Government Printing Office.
- Wolfner, Glenn D., and Richard J. Gelles. 1993. "A Profile of Violence toward Children: A National Study." *Child Abuse and Neglect* 17 (2): 197-212.
- Wulczyn, Fred. 2003. "Closing the Gap: Are Changing Exit Patterns Reducing the Time African American Children Spend in Foster Care Relative to Caucasian Children?" *Children and Youth Services Review* 25 (5-6): 431-62.
- Zellman, Gail L. 1992. "The Impact of Case Characteristics on Child Abuse Reporting Decisions." *Child Abuse and Neglect* 16 (1): 57-74.

Notes

The research in this article was supported through National Institute of Child Health and Human Development grants HD41141-01, HD39135-01, and HD044302-02. The authors are listed in alphabetical order. We thank Jeanne Brooks-Gunn, Jane Waldfogel, and seminar participants at the Johns Hopkins University, Columbia University, Princeton University, Chapin Hall Center for Children at the University of Chicago, and the Institute for Research on Poverty at the University of Wisconsin-Madison for valuable comments.

1. In 2003, 9.8 percent of children classified as non-Hispanic white and 34.1 percent of children classified as black lived in families with income below the poverty line (U.S. Census Bureau 2004).

2. These individuals are not hired specifically to uncover child maltreatment, but in most states, professionals who interact with children (e.g., medical professionals, teachers, social workers, and law enforcement officials) are mandated reporters who are legally required to report suspected maltreatment. In other states, all adults are mandated reporters.

3. Not all cases reported to CPS are investigated. The term "screened-in" is used to denote those cases that are found to be worthy of investigation.

4. In practice, the process is often more complicated. For example, CPS may choose to close the case without an investigation. In some states, cases can either be indicated or substantiated. The former designation is used if there is reason to suspect that maltreatment has occurred, but the allegation cannot be substantiated to the level of evidence required by state law.

5. There is a large economics literature on statistical discrimination. This literature hypothesizes that incomplete information may play a role in race and gender discrimination (Phelps 1972; Aigner and Cain 1977). The theory of statistical discrimination has been used to study racial bias in hiring (Prewett-Livingston et al. 1996; Frazer and Wiersma 2001; Bertrand and Mullainathan 2004), the provision of medical care (Balsa and McGuire 2001, 2003), and even prices charged to consumers for new cars (Goldberg 1996). A related literature in social psychology views discrimination and stereotyping as a result of cognitive processes that rely on racial categorizations when processing information (for a review, see Fiske 1998). A growing economics literature also uses models of categorization to study discrimination (e.g., Fryer and Jackson 2003).

6. The estimates reported here use observations for both marital and nonmarital births and are unweighted. The subsample of nonmarital births in the FFCWS data is nationally

representative of nonmarital births in U.S. cities with populations of 200,000 or more in 1999 (Reichman et al. 2001). Unweighted and weighted estimates using this subsample yield results that are qualitatively similar to those presented here. However, the smaller sample (of 980 rather than 1,417 observations) generally results in less precise estimates.

7. Survey instruments and documentation for the FFCWS can be found at <http://crcw.princeton.edu/fragilefamilies/surveys.asp>.

8. A total of 2,182 mothers completed the full 36-month in-home module with child observations. Of these, 1,582 were either classified as non-Hispanic black or non-Hispanic white. An additional 165 cases were dropped because of missing information on key variables, such as race of the interviewer.

9. Psychometric properties for the CTS-based measures are presented in a study by Straus et al. (1998). The study reports reliability coefficients ranging from 0.55 to 0.70 for the subscales used here.

10. In general, the existing research suggests that the psychometric properties of the HOME are relatively consistent across whites, blacks, and Hispanics (Sugland et al. 1995) and that HOME subscales moderately correlate with other scales measuring similar constructs (Berlin et al. 1995). However, there are also differences in individual item and overall scores across racial and ethnic groups and socioeconomic classes (Bradley et al. 2001). Finally, while the HOME scales tend to predict cognitive and behavioral outcomes, the strength of these predictions varies by race and ethnic group (Berlin et al. 1995; Sugland et al. 1995).

11. The development and psychometric properties of these particular subscales are described in Fuligni, Han, and Brooks-Gunn (2004) and Leventhal, Martin, and Brooks-Gunn (2004). Those studies report that reliability coefficients for lack of warmth range from 0.60 to 0.75 and that coefficients for harshness range from 0.57 to 0.97 across several national samples of children between 2 and 5 years of age. These subscales are sometimes described as measuring punitiveness or hostility and lack of responsivity.

12. This subscale would ideally include a seventh item: whether the mother responded positively when the interviewer praised the child (interviewers were instructed to praise the child at some point in the visit). However, because a large proportion of the data were missing for this item, it is excluded from the subscale.

13. These patterns for the CTS and HOME scales are consistent with results from other studies. For example, there is some evidence that blacks and whites differ in their acceptance of physical discipline and corporal punishment (Straus and Mathur 1996), as well as in their rates of utilizing such practices (Wolfner and Gelles 1993; Dietz 2000). Additionally, regarding racial and ethnic differences in HOME scores, Bradley and colleagues (2001) find that, while there are considerable differences among whites, blacks, and Hispanics, the majority of these differences can be explained by poverty status, something that we will examine in more detail below. Note, however, that there is also evidence that HOME scores have better predictive ability for whites than for blacks or Hispanics in regard to child outcomes (Sugland et al. 1995). Berlin and associates (1995) find that the HOME subscales are better predictors of cognitive and behavioral outcomes for white than for black children. They interpret these findings as evidence that different domains of parenting play a different role for black children than for their white counterparts.

14. Adjustment of income for family size was accomplished by dividing income by the number of adult equivalents in the household. The number of adult equivalents is calculated as the number of adults plus one-half the number of children.

15. The relatively few Hispanic interviewers are classified as white.

16. As noted above, Spanish-speaking Hispanic respondents were paired with Spanish-speaking interviewers. For this reason, Hispanic mothers were excluded from the sample.

17. The ordinary least squares regressions are used because the models are later altered to include interviewer fixed effects and because OLS coefficients represent marginal effects of the predictors on the outcomes, thereby lending themselves to direct interpretation. Results from models used to estimate fixed effects regressions for dichotomous outcome variables (e.g., conditional logit models) cannot easily be converted to marginal effects (although odds ratios can easily be calculated). Results from logit and conditional logit models (not shown) are substantively similar to those presented here.

18. This can be assessed by comparing the coefficients for mother is black across the three models for each outcome.

688 Social Service Review

19. The difference-in-difference estimates for harshness are 8.0 percentage points at the seventy-fifth percentile and 8.8 percentage points at the ninetieth percentile.

20. The CTS are shown to be correlated with other measures of abuse (see Straus et al. 1998), and Berger and Brooks-Gunn (2005) find links between HOME scores and other measures of child maltreatment

Copyright of *Social Service Review* is the property of University of Chicago Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.