

## Data and Methods Used to Construct the Racial and Ethnic Achievement Gap Figures

Sean F. Reardon

Version: February 2, 2015

### Data and Measures

Achievement Data. The data used to compute achievement gaps come from the National Assessment of Education Progress (NAEP). NAEP has two components: the “Long-Term Trend NAEP” (NAEP-LTT) and “Main NAEP.” The NAEP-LTT have been administered to large, nationally-representative samples of 9-, 13-, and 17-year-olds in public and private schools since the early 1970s. NAEP-LTT can be used to provide national trend data, but does not provide state-level information. The Main NAEP state assessments have been administered biennially to state-representative samples of 4<sup>th</sup> and 8<sup>th</sup> graders in every state since 2003, and in many states since the early 1990s. Details about the NAEP assessment program and the content of the assessments can be found on the NAEP website at <http://nces.ed.gov/nationsreportcard/>.

Achievement Gaps. In most cases, we compute achievement gaps from NAEP micro-data. The NAEP micro-data are available through a restricted-access data license from the National Center of Education Statistics (NCES). We measure achievement gaps here using the  $V$ -statistic (Ho, 2009; Ho & Haertel, 2006; Ho & Reardon, 2012). The  $V$ -statistic is a measure of the difference between two distributions. It is very similar to more commonly-used measures of achievement gaps, such as the standardized mean difference between two distributions, but it has the advantage of being independent of the metric in which test scores are measured—it relies only on the ordered nature of test scores, not their cardinal values.

Specifically,  $V$  is defined as follows: let  $P_{w>b}$  indicate the probability that a randomly chosen white student has a higher score than a randomly chosen black student. This will range from 0 to 1, where 0 indicates that no white student has a score as high as the lowest-scoring black student, and 1 means the opposite. A value of 0.5 indicate that there is no achievement gap.  $V$  is defined as a monotonic transformation of  $P_{w>b}$ :  $V = \sqrt{2}\Phi^{-1}(P_{w>b})$ , where  $\Phi^{-1}$  is the probit function. The advantage of this transformation is that it renders  $P_{w>b}$  into a familiar metric:  $V = 0$  implies there is no achievement gap; positive values of  $V$  indicate that white students, on average, score higher than black students; and vice versa. Moreover,  $V$  can be interpreted as a standardized effect size: when the white and black score distributions are both normal,  $V$  will be equal to their difference in means divided by their pooled standard deviation.

For each student in the NAEP micro-data, five plausible test-score values are provided (reflecting measurement error in the tests). To compute a given gap (e.g., in a given state-year-subject-grade), we first compute five non-parametric estimates of  $V$  (based on each of the five plausible values). For each of these estimates, we obtain a standard error by regressing test scores on a racial indicator variable (indicating if a student is black or Hispanic) and taking the standard error from the regression results as the standard error of  $V$  (although there are analytic methods of computing standard errors of non-parametric estimates of  $V$ , it is not clear they work with weighted samples; however, Reardon and Ho (forthcoming) show that when test score distributions are normal—as NAEP distributions generally are—

the standard errors from regression-based gap estimates are very similar to those from non-parametric  $V$  estimates). We then combine the five estimates (Rubin, 1987) to obtain estimates of  $V$  and its standard error.

In some cases, we do not have access to NAEP micro-data. Specifically, the 2013 micro-data are not yet available from NCES. For 2013, we compute gaps from publicly available NAEP data available at <http://nces.ed.gov/nationsreportcard/naepdata/>. We compute the difference in mean scores, divided by the pooled standard deviation of scores (the standardized mean differences). Because the NAEP score distributions are very close to normal distributions, these are very similar to what we would get if we computed  $V$  from the micro-data.

We do not report achievement gaps based unless there are at least 20 students of each group in the sample. In the 2013 data, where NAEP gaps are based on publicly reported data, we report gaps only when the public data includes means and standard deviations for both groups.

State Socioeconomic Disparities Data. The data used to compute state-level measures of racial socioeconomic disparities (income gaps, education gaps, poverty ratios, unemployment ratios, and the summary index of racial socioeconomic disparities) come from the Current Population Survey (CPS) (<http://www.census.gov/cps/>), which collects information on household characteristics, employment, and income for state-representative samples on an ongoing (monthly) basis. We use CPS micro data, obtained from the IPUMS CPS site (<https://cps.ipums.org/cps/>), to compute measures of racial socioeconomic disparities in each state in each year from 2003-2013.

State Socioeconomic Disparities Measures. To compute measures of racial socioeconomic disparities, we first limit the CPS sample to households with children 0-17 years old, so that our measures reflect the socioeconomic disparities among families with children. Because the samples of households in some race/ethnic groups are small in some states, we pool multiple years of CPS data to compute race-specific measures in some cases. If there are at least 100 members of a race/ethnic group in the state sample in a given year, we use data from that year alone. If there are fewer than 100, we pool over 3, 5, or 7 years (using the smallest number of years that yields a sample of at least 100 households) to compute an annual estimate. That is, the annual estimates of average race-specific socioeconomic characteristics are running averages within each state, where the number of years used in the averages depends on the sample sizes within each state.

*Income gap.* Before comparing the income distribution of each racial group, we first convert all incomes to 2013 constant dollars using the Bureau of Labor and Statistics' CPI calculator, found here: <http://data.bls.gov/cgi-bin/cpicalc.pl>. We convert income to constant dollars because in some states the black or Hispanic sample is averaged over multiple years while the white sample is drawn from a single year. We compute income gaps using the  $V$ -statistic described above.

*Parent Education gap.* For each household with children, we construct a variable indicating the highest grade or level of education completed by an adult in the household. We compute education gaps between groups by computing the  $V$ -statistic, which measures the difference between the white and black or white and Hispanic households' educational attainment distributions.

*Unemployment ratio.* For each racial group, we compute the proportion of households with children that have one or more unemployed adults (defined as adults in the labor force but unemployed) in the home. We compute the black/white or Hispanic/white ratio of unemployment rates.

*Poverty ratio.* For each racial group, we compute the proportion of households with incomes that place them below the poverty line, given the size of the household. We compute the black/white or Hispanic/white ratio of poverty rates.

*Racial Socioeconomic Disparities Index.* This is a weighted average of the four socioeconomic disparities measures, where the weights are the coefficients from a regression model. See below (under Figure 5) for more detail.

### Specific Information for Each Figure

Figure 1: “Over the past 40 years, white-black and white-Hispanic achievement gaps have been declining, albeit unsteadily.”

The gaps here are computed from NAEP-LTT data. We fit a precision-weighted least squares cubic regression line through estimated gap trend data, using the inverse of each gap estimate’s squared standard error as weights.

Figure 4: “In the last two decades, racial achievement gaps have been narrowing in most states, but have widened in a small number of states.”

The trends here are precision-weighted least squares cubic, quadratic, or linear fitted trend lines. We show either the cubic, quadratic, or linear trend depending on which fits the data better.

Figure 5: “State racial achievement gaps are strongly correlated with state racial socioeconomic disparities.”

The achievement gaps reported here (on the vertical axis) are random-effects precision-weighted averages of achievement gaps from 2003-2013. Specifically, we fit the random effects model

$$\hat{G}_{ys} = \gamma_0 + \gamma_1(\text{year} - 2008) + (\mathbf{X}_{ys} - \bar{\mathbf{X}}_s)\boldsymbol{\Gamma} + u_s + e_{ys} + \epsilon_{ys}, \quad (1)$$

where  $\hat{G}_{ys}$  is the estimated gap state  $s$  and year  $y$ ;  $\mathbf{X}_{ys}$  is a vector of four state-by-year values of racial socioeconomic disparities (income gap, parental education gap, poverty ratio, and unemployment ratio; note that we subtract 1 from the poverty and unemployment ratios in this regression model so that values of 0 correspond to the condition of equality for all the variables);  $\bar{\mathbf{X}}_s$  is a vector of the state averages of these four variables over the 2003-2013 period;  $u_s$  is a state-specific random effect;  $e_{ys}$  is a state-by-year error term reflecting idiosyncratic temporal variation in the size of the achievement gap, net of the average linear trend and changes associated with temporal variation in  $\mathbf{X}_{ys}$ ; and  $\epsilon_{ys}$  is a state-by-year sampling error term with variance equal to the estimated sampling variance of  $\hat{G}_{ys}$ . After fitting this model, we use  $\hat{G}_s^* = \hat{\gamma}_0 + \hat{u}_s^*$  (where  $\hat{u}_s^*$  is the empirical Bayes estimate of  $u_s$ ) as an estimate of the average achievement gap in state  $s$  over the period 2003-2013. This is plotted on the vertical axis in the figures.

The measures of racial income, education, poverty, and unemployment disparities (plotted on the horizontal axis) are computed for each year 2003-2013 as described above, and then averaged across the years 2003-2013 to compute an average value for each disparity in each state over that time period. When plotted, the poverty and unemployment ratios are in their original metric (one is not subtracted from them) so that they are easily interpretable.

These four measures are combined into a summary index of socioeconomic disparities by adding them to model (1) above:

$$\hat{G}_{ys} = \gamma_0 + \gamma_1(\text{year} - 2008) + (\mathbf{X}_{ys} - \bar{\mathbf{X}}_s)\boldsymbol{\Gamma} + \bar{\mathbf{X}}_s\mathbf{B} + u_s + e_{ys} + \epsilon_{ys}. \quad (2)$$

The predicted value of  $G_s$  in a given state is then  $\hat{\gamma}_0 + \bar{\mathbf{X}}_s\hat{\mathbf{B}}$ . Because  $\bar{\mathbf{X}}_s = \mathbf{0}$  under racial socioeconomic equality,  $\hat{\gamma}_0$  is an estimate of the size of the achievement gap in a state with no racial socioeconomic inequality. We use  $\bar{\mathbf{X}}_s\hat{\mathbf{B}}$ , the part of the average 2003-2013 achievement gap that can be predicted based on the state's racial socioeconomic disparities, as a summary measure of racial socioeconomic disparities. Note that, under this definition, the racial socioeconomic disparities index varies within a state slightly across grades and test subjects, because  $\hat{\mathbf{B}}$  is estimated separately for each grade and subject. When the racial socioeconomic disparities index is used on the x-axis in the figure, the line plotted has an intercept of  $\hat{\gamma}_0$  and a slope of 1: it plots the line  $G = \hat{\gamma}_0 + \bar{\mathbf{X}}_s\hat{\mathbf{B}}$ , which describes the predicted gap for a given value of racial socioeconomic disparities.

Figure 6: "Some states' achievement gaps are higher or lower than would be predicted by their socioeconomic conditions alone."

The figure shows, for each state-subject-grade combination, the difference between the estimated average gap over 2003-2013 (estimated from Model 1 above) and the gap predicted from racial socioeconomic disparities (estimated from Model 2 above). That is, it shows  $\hat{r}_s = \hat{G}_s^* - [\hat{\gamma}_0 + \bar{\mathbf{X}}_s\hat{\mathbf{B}}]$ .

## References

- Ho, A. D. (2009). A nonparametric framework for comparing trends and gaps across tests. *Journal of Educational and Behavioral Statistics*, 34, 201-228.
- Ho, A. D., & Haertel, E. H. (2006). *Metric-Free Measures of Test Score Trends and Gaps with Policy-Relevant Examples*. Los Angeles, CA: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies.
- Ho, A. D., & Reardon, S. F. (2012). Estimating Achievement Gaps From Test Scores Reported in Ordinal 'Proficiency' Categories. *Journal of Educational and Behavioral Statistics*, 37(4), 489-517.
- Reardon, S. F., & Ho, A. D. (forthcoming). Practical Issues in Estimating Achievement Gaps From Coarsened Data. *Journal of Educational and Behavioral Statistics*.
- Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley and Sons.