Available online at www.sciencedirect.com
ScienceDirect
Intelligence 34 (2006) 607–619
INTELLIGENCE

Estimating state IQ: Measurement challenges and preliminary correlates

Michael A. McDaniel

Virginia Commonwealth University, United States

Received 12 March 2006; received in revised form 5 August 2006; accepted 21 August 2006
Available online 10 October 2006

Abstract

The purpose of this study is threefold. First, an estimate of state IQ is derived and its strengths and limitations are considered. To that end, an indicator of downward bias in estimating state IQ is provided. Two preliminary causal models are offered that predict state IQ. These models were found to be highly predictive of state IQ, yielding multiple R's of 0.83 and 0.89. Second, the extent to which state IQ predicts state outcome variables (e.g., gross state product, health, violent crime, and government effectiveness) is estimated. State IQ shows positive correlations with gross state product, health, and government effectiveness and negative correlations with violent crime. These results are consistent with the extent to which IQ predicts outcomes at the level of the individual. Third, a research agenda is provided for improving estimates of state IQ, identifying factors that cause differences in state IQ, and delineating the role of IQ in predicting important variables.

Keywords: National Assessment of Educational Progress (NAEP); Intelligence; IQ; State IQ; Lynn and Vanhanen (2002); Gross state product; Health; Crime; Government effectiveness; Supressor

Lynn and Vanhanen (2002) have estimated the IQ of countries using data from various sources. These IQ estimates have proven useful in the prediction of national wealth. This paper seeks to offer an estimate of state IQ and to discuss strengths and limitations of the estimate. In addition, the paper seeks to evaluate the extent to which estimated state IQ has relationships with other state variables from the disciplines of economics, public health, criminal justice, and political science. Finally, the paper offers a research agenda for improving the estimates of state IQ and for understanding the role of IQ in predicting other important state variables.

State-level data are frequently collected on a wide range of variables and are used for multiple purposes. The U.S. Census Bureau (2006) collects a variety of data that are aggregated at the state level. These state data are used for a variety of government policy purposes including allocation of congressional House seats across states and the targeting of funds for services to various segments of the population. This is a small sampling of the uses for these data. The U.S. Centers for Disease Control (2006) summarizes state-level data on a variety of health issues including trends in health status, care utilization resources, and expenditures (National Center for Health Statistics, 2005). The U.S. Department of Commerce Bureau of Economic Analysis (2006) reports gross state product, an indicator of state productivity. The National Center for Educational Statistics (NCES) (2006) collects and provides state summaries of various educational statistics including measures of cognitive

E-mail address: mamcdani@vcu.edu.

0160-2896/$ - see front matter
doi:10.1016/j.intell.2006.08.007

ability, such as reading and math. The U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics (2006) reports state-level crime data. Non-federal organizations also gather state-level data. The Government Performance Project (2006), a privately funded organization, assesses the effectiveness of state governments. United Health Foundation (2006), a nonprofit private foundation, issues an indicator of overall health status by state. In sum, state-level data are available on a variety of issues.

Little research has addressed the topic of state intelligence. An estimate of the intelligence (IQ) of a state would involve aggregating the IQ of its citizens and assigning the mean of the individuals' IQ as the state IQ. Lynn and Vanhanen (2002) estimated the IQ of countries using data from the average of various measures of intelligence. Their IQ estimates were criticized because of concerns about the representativeness of the samples and the appropriateness of the measures (Barnett & Williams, 2004; Ervik, 2003; Richards, 2002; Volken, 2003). Thus, any effort to develop an estimate of state IQ should specifically address the representativeness of the samples and the appropriateness of the IQ measure. It is ironic that little research has examined the IQ of states given that IQ or IQ-related data are collected routinely in states. For example, college and graduate school entrance exam data are available by state. Likewise, a variety of cognitive ability and achievement tests are collected in primary and secondary schools and are potentially available. Finally, various tests are normed on nationally representative samples and their data could be reported by state.

State IQ can be expected to have correlates with state-level variables given the correlations between IQ and other variables when examined at the level of the individual. For example, IQ is a predictor of productivity at the individual level (Schmidt & Hunter, 1998), thus state IQ should predict productivity at the state level as measured by gross state product. Likewise, IQ predicts health for individuals (Batty & Deary, 2005; Hart,Taylor, & Smith, 2005), thus state IQ should predict state health statistics. IQ also predicts who is likely to engage in crime (Farrington, 2005; Gordon, 1987), thus state IQ should predict state crime data. Intelligent individuals tend to evaluate cognitively complex information more efficiently and accurately, thus making more informed decisions (Gottfredson, 2004; Hunt, 1995). Thus, state IQ may have correlates with the effectiveness of government based on decisions made by elected individuals.

The purpose of this study is threefold. First, an estimate of state IQ is derived and its strengths and limitations are delineated. As part of describing the limitations of the measure, an indicator of downward bias in the estimate of state IQ is developed. Two preliminary causal models that predict state IQ are offered. Second, this paper evaluates whether estimated state IQ is predictive of four state-level variables: gross state product, health, violent crime, and government effectiveness. Third, a research agenda for the study of state IQ is described.

1. Method

1.1. Measures

1.1.1. Estimated state IQ

State IQ was estimated from the National Assessment of Educational Progress (NAEP) standardized tests for reading and math that are administered to a sample of public school children in each of the 50 states. There is a substantial technical literature on these measures (NCES, 2006). Samples of the NAEP items are presented in Appendix A. State data were available for grades 4 and 8. Twelfth grade data are collected but are not available by state and thus were not used. NAEP data are available for multiple years. Not all states participated in each administration of each test in each year, but all states participated in the tests for multiple years. For each year, for each test, the national mean and standard deviation was used to standardize the test to have a mean of 100 and a standard deviation of 15. This standardization places the scores on the typical metric for IQ tests. The means of the standardized reading scores for grades 4 and 8 were averaged across years as were the means of the standardized math scores. State IQ was defined as the average of mean reading and mean math scores.

1.1.2. Percent of white children in non-public schools

Only public school NAEP data are published by state. State IQ scores derived from the NAEP tests will not be representative of the states' populations to the extent that the public school children who participate in the testing are not representative of the states' populations. Some children do not attend public schools. They might be educated at home, they might have private tutors, or they might attend private schools. Most private schools require the child's parents to pay tuition and fees unlike public schools, which are funded by the government. Average income differences between Whites, Blacks, and Hispanics make it more likely that a larger percent of Whites than Blacks or Hispanics will attend non-public schools. On average, it is expected that those who do not attend public schools are more cognitively gifted than those who attend public schools for several reasons. First, some private schools have cognitively-based selection criteria. Second, parents who remove their children from public schools, particularly bad public schools, may have a greater than average interest in their children's education. Third, there is a positive correlation between IQ and income. Thus, cognitively-gifted students of all racial groups are more likely to attend private schools. Given the large mean cognitive ability differences between Whites, Blacks, and Hispanics (Roth, BeVier, Bobko, Switzer, & Tyler, 2001), one would expect a larger percentage of White students to be in private schools than Black or Hispanic students. Finally, when a public school has primarily Black or Hispanic students, some White parents may choose to send their children to non-public schools due to attitudes toward diversity, perceptions of safety, or social and academic concerns. Thus, for several reasons, it is anticipated that White and/or cognitively-gifted students will be more likely than others to attend non-public schools.

It is argued that state IQ scores derived from public school data will underestimate state IQ to the extent that the percent of White students in the public schools in the state is less than the percent of Whites in the state. Thus, this disparity was estimated using year 2000 school data from the NAEP data set and year 2000 population data from the census data. It is asserted that as the percent of White children in non-public schools increases, the estimated state IQ based on public school data will underestimate the population state IQ (the true value of the population state IQ). In this definition, children in non-public schools would include children in private schools, home schooling, or any schooling outside of public schools.

Tzelgov and Henik (1991) addressed how past researchers have sought to improve prediction by removing systematic measurement error such as halo or leniency effects in ratings. In the performance appraisal literature, a severity measurement error is the opposite of a leniency error, in that the error results in a systematic underestimate of the true value. By way of metaphor, one can view the percent of White children in non-public schools as an indicator of severity measurement error such that as the indicator increases, the downward bias in estimated state IQ increases. It is suspected that in a regression context, the percent of White children in non-public schools combined with estimated state IQ will result in a suppression situation (Tzelgov & Henik, 1991). Specifically, it is suspected that the standardized beta weight for estimated state IQ will be larger when the percent of White children in non-public schools is added as a second predictor than when estimated state IQ is the sole predictor. Expressed another way, it is anticipated that the zero-order correlation (which is the same as the standardized beta weight with one independent variable) between estimated state IQ and other state variables will be an underestimate of the actual predictive value of estimated state IQ. This suppression effect, the increase in the beta weight of estimated state IQ, is expected because the severity bias in the estimated state IQ measure will be reduced by the addition of the percent of White children in non-public schools to the regression equation.

1.1.3. Pupil/teacher ratio

The pupil/teacher ratio is a measure of class size and was obtained from NCES (2006).

1.1.4. Expenditure per student

This dollar value is the state expenditure per student and was obtained from NCES (2006).

1.1.5. Percent Black in the state, percent Asian in the state, percent Hispanic in the state

The first two variables are the percent of the state census respondents who identified their race as Black or Asian in the 2000 census as reported by the U.S. Census Bureau (2006). Hispanic is not a race in the census data. The percent Hispanic is the percent who indicated that their ancestry was Hispanic regardless of their race.

1.1.6. Low birth weight

This variable, obtained from the Center for Disease Control (2006), is the prevalence of low birth weight as measured by the percentage of live born infants weighing under 2500 g at birth. The variable is the mean of the data from the years 1994 to 1996.

1.1.7. Percent receiving no prenatal care

This variable obtained from the Center for Disease Control (2006) is the percent of mothers delivering live infants who did not receive care during the first trimester of pregnancy. The variable is the mean of the data from the years 1994 to 1996.

1.1.8. Gross state product

As a measure of state productivity, the total gross state product per capita averaged across years 2000 to 2004 was used. The gross state product is reported by the U.S. Department of Commerce (2006). To calculate the per capita version of this variable, it was divided by the state population as reported in census data. In addition to the decennial census, the U.S. Census Bureau collects samples of state data every year permitting estimates of state population for each year. These yearly census data permitted the calculation of the per capita values of gross state product for the years 2000 to 2004.

1.1.9. State health

As a measure of state health, the state health scores were averaged for the years 2002 through 2005, as reported by United Health Foundation (2006; per request of the United Health Foundation of the data, the 2005 data are to be cited as "America's Health Rankings-2005 Edition, © 2005 United Health Foundation. All Rights Reserved."). This health measure is an aggregate of various indices of the health of the states' residents.

1.1.10. Violent crime

This measure is drawn from the Bureau of Justice Statistics (2006) and reports violent crimes per thousand people. It is the mean of the years 2002 through 2004.

1.1.11. Government effectiveness

The Government Performance Project (2006) is a privately funded organization that rates the effectiveness of state governments in several areas. Data were available for 1991, 2001, and 2005. Government effectiveness is defined in this study as the mean of the overall effectiveness rating for the 3 years.

1.1.12. Analyses

The reliability of the measures were based on data aggregated across tests (for the NAEP data) or across years (for the NAEP data and the gross state product, health, violent crime, government effectiveness). The relationship between state IQ and the other variables using correlation and multiple regression was examined. The multiple regression analyses were used to evaluate the presence and strength of suppression situations in regressions predicting gross state product, health, violent crime, and government effectiveness.

2. Results

2.1. Descriptive statistics

Table 1 presents means and standard deviations for all variables as well as a full correlation matrix.

2.2. Reliability of measures

Alpha reliability was assessed for all variables that were composites of the same variable across years. Table 1 presents the reliability when available as the diagonal of the correlation matrix. The reliability of state IQ was estimated by calculating a coefficient alpha for 6 years of 4th grade reading data (1992, 1994, 1998, 2002, 2003, 2005), 4 years of 8th grade reading exams (1998, 2002, 2003, 2005), 4 years of 4th grade math (1992, 2000, 2002, 2005) and 5 years of 8th grade math (1990, 1992, 2000, 2003, 2005). Sixteen states had data for all tests for all years and those data yielded an alpha reliability of 0.99. The alpha based on data from all states required that some test intercorrelations be based on more states than others. That alpha reliability was also 0.99. Alpha reliabilities were calculated for percent of underweight babies (alpha = 0.88), percent of mothers who did not receive prenatal care in the first trimester of pregnancy (alpha = 0.96), gross state product (alpha = 0.99), health (alpha = 0.99), violent crime (alpha = 0.99) and government effectiveness (alpha = 0.92). The extremely high reliabilities obtained do not rule out measurement error in a state's data for a given year or in the composites of the yearly data. However, the reliabilities do indicate that the rank order of states on these variables is remarkably stable. Thus, although some states may improve their reading and math test scores, reduce their crime, or alter the resident's health status across years, such efforts have not appreciably changed the rank order of the states on these variables.

2.3. Prediction of estimated state IQ

Table 2 (Model 1) shows the regression results for the prediction of estimated state IQ. Using percent of Blacks, percent of Hispanics, percent of Asians, pupil/teacher ratio, and state expenditures per student, a multiple R of 0.83 was obtained ($F = 19.80$, $df = 5,44$, $p < 0.0001$). States with higher estimated state IQ have a smaller proportion of Black, Hispanic, and Asian residents, larger expenditures per student and smaller class sizes. One can alter the magnitude of the standardized beta weights for some racial groups by adding health variables that are highly correlated with race. A state's percent of underweight babies is highly correlated with the percent of the state which is Black ($r = 0.79$). A state's percent of mothers who did not receive prenatal care in the first trimester of pregnancy is strongly correlated with the percent of Hispanics in the state ($r = 0.62$). Table 2 (Model 2) shows the results when these variables are added ($R = 0.88$, $F = 22.82$, $df = 7,42$, $p < 0.0001$). The addition of the health variables reduces the beta weights for the percent Black and the percent Hispanic in the states.

Both models provide excellent prediction of estimated state IQ (multiple R's of 0.83 and 0.88). In Table 2, the states that are not well predicted by the regression given their large negative or positive residuals are provided.

Table 1
Means, standard deviations and correlation matrix of all variables

Variables	Mean	σ	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Estimated state IQ	110.34	2.70	0.99												
2. % Whites not in public school	5.90	3.92	−0.63	–											
3. Pupil/Teacher ratio	15.67	2.14	−0.38	0.17	–										
4. Expenditure per student	8206.03	1522.88	0.39	−0.03	−0.32	–									
5. % Black	9.73	9.53	−0.51	0.69	−0.06	−0.05	–								
6. % Asian	2.88	5.77	−0.27	−0.04	0.20	0.07	−0.11	–							
7. % Hispanic	7.79	8.91	−0.34	0.40	0.36	0.01	−0.16	0.14	–						
8. % Low birth weight	7.30	1.20	−0.71	0.66	0.03	−0.15	0.79	−0.02	0.09	0.88					
9. % No prenatal care	18.40	4.20	−0.58	0.47	0.30	−0.30	0.13	−0.02	0.62	0.36	0.96				
10. Gross state product	0.035	0.007	0.28	0.14	−0.01	0.69	−0.03	0.18	0.14	−0.07	−0.26	0.99			
11. State health	2.51	11.29	0.75	−0.62	−0.06	0.35	−0.70	0.21	−0.07	−0.79	−0.54	0.34	0.99		
12. Violent crime	0.004	0.002	−0.58	0.75	0.28	0.00	0.54	−0.02	0.41	0.61	0.45	0.14	−0.66	0.99	
13. Government effectiveness	4.86	1.50	0.34	−0.10	0.10	0.00	0.12	−0.21	−0.21	−0.13	−0.25	0.10	0.12	0.02	0.92

The diagonal contains alpha reliabilities across years indicating the stability of the rank order of the states.

The author does not have any compelling explanation for these states having large residuals, but they are provided in hopes of generating ideas for why these state IQs are not well predicted by the regression equations.

Both models showed negative beta weights for percent of the state that is Asian. The percent Asian variable has a near zero correlation with the two health variables and thus the inclusion of the health variables does little to alter the beta weights for percent Asian. The negative beta weights in both models for Asians are contrary to the research showing that Asians tend to have higher mean IQs than other racial/ethnic groups. Given the relatively small percent of Asians in any given state, the effect may not be large enough to affect the estimated state IQ.

2.4. Prediction of gross state product, health, violent crime, and government effectiveness

Table 3 lists the variables analyzed for each of the 50 states. These results are provided so that they can be replicated or extended by interested parties. Table 4 presents the results of a series of regression equations in which the independent variables are state IQ and the percent of White children in non-public schools.

Table 2
Causal model statistics

Predictor	Standardized B[a]	Standard error	95% Confidence interval[b]		States with large residuals (1.5 or greater)	
			Lower	Upper	Negative	Positive
Model 1						
% Black	−0.59	0.085	−0.77	−0.42	Alaska	Texas
% Hispanic	−0.36	0.091	−0.54	−0.18	New Mexico	Virginia
% Asian	−0.29	0.087	−0.46	−0.12	Rhode Island	
Expenditure per student	0.35	0.090	0.17	0.54	West Virginia	
Pupil teacher ratio	−0.11	0.097	−0.31	0.08		
Model 2						
Low birth weight	−0.43	0.132	−0.69	−0.16	Alabama	Colorado
No prenatal care first trimester	−0.23	0.105	−0.45	−0.03	Alaska	
% Black	−0.19	0.128	−0.45	0.07	New Mexico	
% Hispanic	−0.10	0.104	−0.30	0.11	Rhode Island	
% Asian	−0.28	0.074	−0.42	−0.13	West Virginia	
Expenditure per student	0.23	0.082	0.06	0.39		
Pupil teacher ratio	−0.14	0.082	−0.30	0.03		

[a] The standardized beta weights are from a simultaneous regression.
[b] When the confidence intervals do not contain zero, the standardized beta-weight is statistically significantly different from zero ($p<0.05$).

Table 3
Listing of study data

State	Estimated state IQ	% Whites in non-public education	Gross state product	Health	Violent crime	Government effectiveness
Alabama	95.7	9.5	0.028	−11.9	0.004	1.3
Alaska	99.0	6.1	0.047	−0.8	0.006	3.3
Arizona	97.4	11.0	0.032	−1.1	0.005	4.3
Arkansas	97.5	6.8	0.027	−14.3	0.005	3.0
California	95.5	10.6	0.040	4.9	0.006	2.7
Colorado	101.6	6.2	0.041	12.5	0.004	4.0
Connecticut	103.1	7.4	0.049	15.6	0.003	3.0
Delaware	100.4	11.8	0.059	−2.3	0.006	6.7
Florida	98.4	12.2	0.032	−9.9	0.007	4.7
Georgia	98.0	8.0	0.036	−9.5	0.005	5.0
Hawaii	95.6	2.5	0.036	15.1	0.003	2.7
Idaho	101.4	2.0	0.029	8.6	0.002	4.3
Illinois	99.9	8.1	0.039	0.3	0.006	5.0
Indiana	101.7	2.2	0.034	0.5	0.003	4.3
Iowa	103.2	2.4	0.034	14.5	0.003	6.3
Kansas	102.8	4.4	0.033	7.2	0.004	5.7
Kentucky	99.4	1.7	0.030	−7.8	0.003	6.7
Louisiana	95.3	13.7	0.031	−20.5	0.006	5.3
Maine	103.4	0.0	0.030	14.3	0.001	4.3
Maryland	99.7	8.7	0.037	−0.9	0.007	6.3
Massachusetts	104.3	5.8	0.046	16.9	0.005	4.3
Michigan	100.5	4.7	0.035	0.8	0.005	7.3
Minnesota	103.7	5.2	0.040	23.3	0.003	6.3
Mississippi	94.2	13.4	0.024	−20.7	0.003	4.0
Missouri	101.0	4.5	0.033	−3.3	0.005	7.0
Montana	103.4	3.3	0.026	4.0	0.003	4.3
Nebraska	102.3	4.3	0.036	11.2	0.003	5.7
Nevada	96.5	8.5	0.039	−5.5	0.006	4.0
New Hampshire	104.2	−0.4	0.037	22.2	0.002	3.3
New Jersey	102.8	5.7	0.044	9.0	0.004	5.0
New Mexico	95.7	9.5	0.029	−7.5	0.007	3.3
New York	100.7	7.1	0.043	−0.4	0.005	3.7
North Carolina	100.2	9.2	0.036	−5.6	0.005	5.3
North Dakota	103.8	2.4	0.032	14.5	0.001	5.0
Ohio	101.8	3.3	0.034	1.8	0.003	6.0
Oklahoma	99.3	9.2	0.028	−11.2	0.005	3.3
Oregon	101.2	3.1	0.033	7.9	0.003	4.3
Pennsylvania	101.5	5.8	0.035	3.2	0.004	6.3
Rhode Island	99.5	7.6	0.035	11.6	0.003	3.0
South Carolina	98.4	11.3	0.030	−15.3	0.008	6.3
South Dakota	102.8	1.5	0.034	8.9	0.002	4.7
Tennessee	97.7	6.8	0.033	−13.8	0.007	4.7
Texas	100.0	10.4	0.036	−4.6	0.006	6.0
Utah	101.1	−0.5	0.032	18.2	0.002	8.0
Vermont	103.8	−0.2	0.032	19.8	0.001	5.3
Virginia	101.9	6.5	0.040	6.8	0.003	7.7
Washington	101.9	4.5	0.039	11.5	0.003	7.7
West Virginia	98.7	−0.1	0.025	−10.1	0.003	3.7
Wisconsin	102.9	6.6	0.035	12.7	0.002	5.3
Wyoming	102.4	1.0	0.041	4.3	0.003	3.0

The correlation matrix (Table 1) shows that estimated state IQ has positive correlations with gross state product ($r=0.28$), state health ($r=0.75$), and government effectiveness ($r=0.34$). Estimated state IQ correlated inversely with violent crime ($r=-0.58$). Thus, states with higher estimated state IQ have greater gross state product, citizens with better health, more effective state governments, and less violent crime.

Suppression situations are uncommon in the psychological sciences and can initially seem complex. In the

Table 4

Criterion variables	Beta weight for state IQ as the sole predictor (Beta weight for state IQ when % of White children in non-public schools is in the equation)	Beta weight for % of White children in non-public schools as the sole predictor (Beta weight for % of White children in non-public schools when state IQ is in the equation)	R
Gross state product	0.28 (0.60)	0.14 (0.51)	0.49[a]
Health	0.75 (0.59)	−0.62 (−0.25)	0.78[a]
Violent crime	−0.58 (−0.17)	0.75 (0.65)	0.77[a]
Government effectiveness	0.34 (0.46)	−0.10 (0.19)[b]	0.37[a]

[a] All R's are statistically significant ($p<0.03$).
[b] The signs of these beta-weights (−0.10 and 0.19) are correct. Beta weights sometimes change signs in suppression situations.

text that follows, regression equations are described in which the suppression situation occurs and those in which it is absent. The substantive meaning of why suppression occurs and why it does not occur is explained. Because this is somewhat complex, an overview of the findings is first provided. Then, the results and the corresponding table are presented. Finally, the suppression arguments are revisited.

When the suppression situation occurs, the standardized beta weights in the two predictor regressions are *larger* than the zero-order correlations. This is the empirical definition of reciprocal suppression (Tzelgov & Henik, 1991). The substantive reason why this occurs is because each predictor in the regression partials out *criterion-irrelevant variance* in the other. It will be shown that this criterion-irrelevant variance is related to the racial composition of the state. Thus, race is not substantially predictive of the criterion in the regression where a suppression situation exists. When suppression does not occur, the standardized beta weights in the two predictor regressions are *smaller* than their zero-order correlations. This is the empirical definition of a redundancy situation and is typical of most regression applications (Tzelgov & Henik, 1991). The substantive reason for the redundancy situation is that each predictor in the regression partials out *criterion-relevant variance* from the other. For these regressions, the criterion-relevant variance is also related to the racial composition of the state. Thus, for two criteria in which a redundancy situation exists, race is very predictive of the criteria. In brief, the regressions show a suppression situation when race is not substantially correlated with the criterion and the regressions show a redundancy situation when race is substantially correlated with the criterion.

Table 4 provides the results of the regression analyses that evaluated the presence of a suppression situation where the independent variables are the percent of White children in non-public schools and estimated state IQ. The first column is the criterion variable. The second column contains the standardized beta weights for estimated state IQ for predicting each criterion. The first beta weight is for estimated state IQ as the sole predictor in the equations (that beta weight is equal to the zero-order correlation). In parentheses in column two is the standardized beta weight for state IQ when the percent of White children in non-public schools is added to the equation. To the extent that the second (parenthetical) standardized beta weight is larger than the first standardized beta weight, then a suppression situation exists. Thus, for the prediction of gross state product, the beta weight for estimated state IQ as the sole predictor is 0.28 which increases to 0.60 when the percent of White children in non-public schools is added to the equation. The suppression situation exists for the prediction of gross state product and government effectiveness, but not for the prediction of health and violent crime. The third column contains the standardized beta weights for the percent of White children in non-public schools. The first beta weight is for the regression where percent of White children in non-public schools is the sole predictor (this beta weight is the same as the zero-order correlation). In parentheses is the beta weight for percent of White children in non-public schools when estimated state IQ is a second predictor in the equation. To the extent that the second (parenthetical) beta weight is larger than the first beta weight, evidence of a suppression situation exists. Thus, for predicting gross state product, the beta weight for estimated percent of White children in non-public schools as the sole predictor is 0.14, which increases to 0.51 when estimated state IQ is added to the equation. The regressions predicting gross state product and government effectiveness show a suppression situation. The last column lists the multiple R for the two predictor regression equations.

As described above, two of the four regressions indicate a suppression relationship (Tzelgov & Henik, 1991). For gross state product and effectiveness of state government, the zero-order correlation between estimated state IQ and these variables is substantially lower than its beta weight of estimated state IQ in a two predictor regression that includes percent of White children in non-public schools. Thus, the zero-order correlations between estimated state IQ and these criteria are sharp underestimates. Also, in these regressions, the zero-order correlation between percent of White children in non-public schools is sharply lower than the beta weight of percent of White children in non-public schools in a two predictor regression that includes estimated state IQ. Thus the zero-order correlations between percent of White children in non-public schools and these criteria are sharp underestimates. Since the zero-order correlations of estimated state IQ and percent of White children in non-public schools are each suppressed by the other variable, one would describe this as a reciprocal suppression relationship (Tzelgov & Henik, 1991).

The substantive meaning of this reciprocal suppression relationship is related to the racial composition of the states. In part, the percent of White children in non-public schools is a function of the percent of Blacks in the state. For a variety of reasons, some White parents will remove their children from public schools if the percent of Blacks in public schools is relatively large. This relationship is illustrated by the 0.69 correlation between percent of Blacks in the state and the percent of White children in non-public schools. Thus, both the percent of Blacks in the state and the percent of White children in non-public schools are both indicators of the relative proportions of Blacks and Whites in the states. The criteria whose regressions show reciprocal suppression have relatively low correlations with both percent of the Black population and percent of White children in non-public schools (gross state product, r's = −0.03, 0.14; government effectiveness, r's = 0.12, −0.10). Thus the racial composition of the state is not very relevant to the criteria. In the two predictor regressions, the beta weight for estimated state IQ increases because the criterion-*irrelevant*-race variance is reduced by the variable percent of White children in non-public schools. Also, the beta weight for the percent of White children in non-public schools increases because its criterion-*irrelevant* race variance is reduced by the estimated state IQ variable. Thus, both variables increase their prediction of the criteria.

The relative racial composition of the state also explains the regression results when a suppression situation is not evident (the regressions for the prediction of health and violent crime). These regressions indicate a redundancy situation (Tzelgov & Henik, 1991) and are typical of most regression applications. Both of these variables have substantial correlations with percent of the state which is Black (health, −0.70; violent crime, 0.54) and the percent of White children who are in non-public education (health, −0.62; violent crime, 0.75). Thus, the racial composition of the states is strongly predictive of state health and state violent crime. When both estimated state IQ and percent of White children in non-public schools are predictors in the regression, each variable partials out of the other the criterion-*relevant* race variance and thus the beta weights for each variable are lower than their zero-order correlations.

Often researchers will refer to suppressor variables when a more appropriate description is suppressor situation (Tzelgov & Henik, 1991). These analyses illustrate why suppression situation is a more descriptive term. In the analyses, estimated state IQ and percent of White children in non-public schools combined to form a suppression situation in the prediction of criteria where race was not an important predictor. Thus, one might be tempted, incorrectly, to call one or both of them suppressor variables. However, in two other regressions, in which race was important for predicting criteria, no suppression was evident. Thus, the suppression situation was not due to suppressor variables but to the interrelationships among the variables. When the criteria did not have a substantial relationship with race, a suppression situation occurred. When the criteria did have a substantial relationship with race, a suppression situation did not occur. Because the same variables (estimated state IQ and percent of Whites in non-public schools) are involved in suppression relationships for some but not all criteria, it is best to refer to the instances of suppression as suppressor situations rather than to refer to the two predictors as suppressor variables.

3. Discussion

Most of the measures in this study were aggregates of multiple years of data to reflect the status of states on average. Consider gross state product. In a given year, a state may have an unusually low gross state product due to a natural disaster (e.g., hurricane) or economic conditions (e.g., oil prices or terrorism leading to reduced tourism) that affect one state more than another. Data were aggregated across years which served to balance out these kinds of effects. The reliability data reported are alpha reliabilities of the variable across years. These reliability statistics are best interpreted as an indication of the stability of a variable over time. All the reliabilities were above 0.88. The median reliability was 0.99. These high reliabilities indicate that the rank order of the states was extremely stable across years for any given variable.

3.1. Development and evaluation of a measure of estimated state IQ

The first goal of this paper was to develop and estimate state IQ and examine the strengths and weakness of the measure. Reading and math tests are excellent measures of intelligence (Jensen, 1998). Critics of the

NAEP reading and math data as the basis for an IQ measure may attempt to raise a distinction between achievement and intelligence tests. Specifically, the NAEP reading and math test are academic achievement tests and some might argue that they are either (1) not exactly the same as intelligence tests or (2) nothing like intelligence tests. This is an old debate. Coleman and Cureton (1954) documented the very substantial overlap in content between measures labeled achievement and measures labeled intelligence. Cronbach (1984) noted that the distinction "is one of point of view, more than test content" (p. 32). In more recent times, Frey and Dettermam (2004) have argued that the Scholastic Assessment Test (SAT) is "mainly a test of g" (p. 273). Koenig (2006) has made similar arguments for the American College Test (ACT). Even when one treats achievement and intelligence tests as distinct forms of assessment, they are very highly correlated (Deary, Strand, Smith & Fernandes, in press). Deary et al. (in press) reviewed the literature and concluded that there is substantial agreement that intelligence and achievement are highly correlated. Thus, we offer the state IQ estimate based on reading and math scores as a reasonable measure of state IQ. Those who attempt to define intelligence with less research-based definitions of intelligence (practical intelligence, successful intelligence, emotional intelligence) may criticize the measure, but it is clear that the use of reading and math as measures of general cognitive ability is well supported by the intelligence literature.

The estimated state IQ measure can be criticized because it is estimated from a sample that is not representative of the state's population. Not all students attend public schools and the estimated state IQ is likely to underestimate the population state IQ as the percent of children in non-public schools increases. The regression analyses showed that the percent White children in non-public schools was an indicator of bias in the estimated state IQ measure. It is noted that states may vary in the decision rules for who is exempted from the NAEP data due to cognitive impairments (McLaughlin, Gallagher, & Stancavage, 2004). Pressures placed on states to improve scores in reading and math (which these analyses show to be extremely stable across years), may cause some states to give the impression of improvement by excluding larger percentages of cognitively ill-equipped students in order to inflate their mean test scores. For example, it appears that the state of Texas has systematically distorted test results on the Texas Assessment of Academic Skills primarily through coaching on test content (Klein, Hamilton, McCaffrey, and Stecher, 2000). Coaching improves test score data on the Texas test but does nothing to improve actual reading and math skills as assessed by NAEP exams. Teaching to the content of state tests is likely to be common in all states and Texas should not be singled out. Texas is cited specifically because the extent of the coaching problem is well documented for that state.

The initial path model yielded a multiple correlation of 0.83. States with higher estimated state IQ have a smaller proportion of Black, Hispanic, and Asian residents, more expenditures per student and smaller class sizes. The inclusion of state racial composition as a primary cause of state IQ may be objectionable. However, there are large mean racial differences in IQ (Roth et al., 2001) and these differences manifest themselves in the state-level data. In the second model, the percent of children born with low birth weight and the percent of live births where the mother received no prenatal care in the first trimester of pregnancy were added as predictors. Both variables have substantial negative correlations with estimated state IQ (r's $=-0.71, -0.58$). As seen in Table 1, low birth weight is strongly correlated with the percent of the Blacks in the state ($r=0.79$). No prenatal care in the first trimester of pregnancy is strongly correlated with the percent of the Hispanics in the state ($r=0.62$). Neither variable is related to the percent of Asians in the state (both correlations are -0.02). This pattern of correlations has the potential to reduce the contribution of the percent Black and percent Hispanic for predicting state IQ. As can be seen in Table 2 (Model 2), the addition of these variables substantially reduces the standardized beta weights for the percent Black and Hispanic but does little to the beta weight for percent Asian. Thus, one can reduce the apparent impact of race on state IQ by adding health or other variables strongly correlated with race.

Substantial research can be conducted to improve estimates of state IQ and to identify the factors that cause state IQ differences. Suggestions are presented below for improving state estimates of IQ and for exploring the causes of state IQ.

The measure of state IQ used in this study is highly cognitively loaded, consistent with the traditional view of IQ as a measure of cognitive capacity. The exploration of state differences in other conceptualizations of intelligence and in other important individual difference variables is needed. For example, can the stereotypes concerning the cordiality of Southerners and the lower cordiality of residents New York be confirmed by data? Are residents of Missouri, the "show me" state, more skeptical than others?

More representative estimates of state IQ could likely be obtained. For example, many cognitive ability tests are normed on national samples. Although state data from a nationally representative sample are not necessarily representative by state, it may be possible to estimate representative values from such data. To the extent that estimates of state IQ derived from different data sources and different construction strategies agree, one can have greater confidence in the quality and accuracy of the measure of state IQ.

Research into the causes of state differences in IQ can be informed by knowledge of the reasons or correlates of the individual differences in IQ. At the level of individuals, IQ is clearly a function of both genetics and environment (Bouchard, 2004). Although the specific genetic and environmental mechanisms that cause IQ are relatively unknown, research in these areas is on going. It is reasonable to suggest that the genetic and environmental factors that cause individual differences in IQ, will also cause differences in state IQ.

IQ at the individual level has strong correlates with race. There are large and intractable mean racial differences in IQ at the person level. The differences are termed *intractable* because they have been relatively constant across decades and have not been appreciably affected by environmental interventions (Murray, 2005). Because racial composition of the state is a large magnitude correlate of state IQ, one cannot expect meaningful changes in estimated state IQ as long as state racial composition is relatively stable. While increased education expenditures and smaller class sizes are to be encouraged, the stability of the rank order of NAEP test data suggests that states are not going to alter their standing on estimated state IQ dramatically through such efforts.

3.2. Strategies to raise state IQ

Since state IQ differences are a function of the IQ of state residents, state IQ could change if the residents of the state change. Thus, states might structure incentives to encourage those with high IQs to remain in the state. Likewise, a state may encourage high IQ individuals to have children. Over time, these policies should raise the average IQ of state residents. Below strategies that might be used to raise state IQ are offered.

Parents have increasingly used genetic testing to determine if they will have children. For example, genetic testing has resulted in the substantial reduction of children born with Tay Sachs (Strom, 2004). Down's syndrome is a genetic disorder that results in mental retardation and routine genetic testing may be responsible for the recent decline in the births of children with Down's syndrome (Kristol, 1993). Some genetic testing can be done on the parents. Other genetic testing is done on the fetus and would serve to reduce the birth of cognitively-impaired individuals if the fetuses were aborted. Should genetic tests be developed and routinely administered for conditions associated with low IQ, one might expect some parents to forgo having children or to voluntarily terminate pregnancies likely to result in the birth of a low IQ child.

Genetic testing is not the sole genetic-based approach to enhancing IQ. Parents often devote substantial resources to improving the chances of their children's success. In the future, genetic technologies are likely to be developed to permit parents to select characteristics of their children (Agar, 2004). Some parents may take advantage of these technologies to genetically program enhanced IQ in their children.

Some have suggested government intervention in decisions about who may have children. In reaction to Lykken's (2001) proposal for the state to license those who may have children, Scarr (2000) offered an alternative to licensing by making effective contraception and abortion readily available. Such voluntary programs would need to be evaluated to determine their effect on IQ.

States might alter state IQ by influencing who lives in the state. For example, a state might encourage businesses that rely on highly educated employees to relocate to their state. States might also increase the selectivity of their universities so that they attract higher IQ faculty and students. The influx of higher IQ faculty and their families would directly impact state IQ. High IQ college students might be given incentives to stay in the state after graduation. Although not tied to IQ, the city of Kalamazoo, Michigan has started to provide free or reduced college tuition to city residents who are educated in the public schools (Boudette, 2006). Incentives associated with higher education opportunities are likely to attract families who value education. Such plans may result in attracting higher IQ individuals to the locality. Some ethnic groups such as Ashkenazi Jews and some Asian populations have higher than typical IQs and states may take steps to make their state appealing to individuals in these groups. For example, some colleges make their campuses Jewish-friendly to encourage more Jewish students with high cognitive skills to attend their universities (Wiener, 2002).

Some of these strategies (e.g., attracting those with high IQ to move to the state) might raise the IQ of a state at the expense of IQ in another state. As such, they may not be reasonable policies for the nation as a whole. Other strategies such as genetic testing for IQ-related disorders would benefit a state without harming other states. It is recognized that some of these strategies (readily available contraception, genetic testing coupled with abortion) are likely to be objectionable to some. Others might find it objectionable not to follow strategies to raise the IQ of states, particularly if subsequent research determines that state IQ is a cause of important state objectives such as increasing productivity, health, and government effectiveness or in reduction of crime. These strategies are offered to encourage debate on how states, and the nation as a whole, may take actions to increase IQ. It is hoped that future thought and research can inform policies that can increase state IQ.

3.3. Estimated state IQ in the prediction of important state variables

The second goal of the paper was to address the relationship between estimated state IQ and gross state product, health, violent crime, and government effectiveness. Estimated state IQ was positively related to gross state product, health, and government effectiveness and negatively related to violent crime. All these relationships are consistent with what one would expect based on the correlates of IQ at the level of the individual. These variables span the disciplines of economics, criminal justice, and public administration. The explanations for these correlations may appear naïve and those with greater expertise are encouraged to extend research on the role of state IQ in predicting and/or causing these variables.

3.3.1. Gross state product

Gross state product is substantially correlated with estimated state IQ ($r=0.28$, $B=0.60$). Additional research should focus on subcategories of gross state product. Some areas of gross state product may be more related to state IQ than others. For example, states that are rich in natural resources that can be mined are likely to have larger gross state products due to mining than state's lacking resources regardless of state IQ.

3.3.2. Health

Estimated state IQ is a substantial correlate of the state health variable ($r=0.75$). Additional research should focus on subcategories of state health. Some health conditions may be more correlated with state IQ than others. Because race is also a strong correlate of state health, analyses should examine the joint and unique effects of state IQ and race on health.

3.3.3. Violent crime

Estimated state IQ is a substantial correlate of violent crime ($r=-0.58$). Additional research could focus on subcategories of violent crime or other categories of crime. Some types of crime may be more correlated with state IQ than others. Whereas race is also a strong correlate of violent crime, analyses should examine the joint and unique effects of state IQ and race on violent crime.

3.3.4. Government effectiveness

States with higher estimated IQ have more effective government ($r=0.34$; $B=0.46$). This may occur because more intelligent individuals tend to vote for leaders who can effectively manage and direct large organizations. Another possibility is that higher IQ states have fewer problems with respect to health and crime and greater gross state product. It is likely easier to manage a state when there are fewer problems and greater productivity than in a state facing many problems with limited productivity. Government effectiveness can be measured in ways other than an overall effectiveness measure. For example, for the year 2005, ratings are also available for government effectiveness in the areas of money, people, infrastructure, and information (http://results.gpponline.org/).

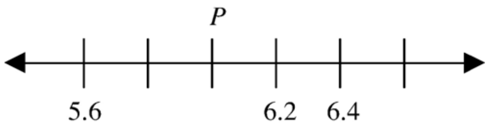
Our correlational analyses are limited. Certainly, more careful reasoning and causal models are needed to understand fully the relationship between estimated state IQ and gross state product, health, violent crime, and government effectiveness.

3.4. Comparison to Lynn and Vanhanen (2002)

Lynn and Vanhanen's (2002) research on national IQ was a substantial effort that inspired the current study. However the Lynn and Vanhanen efforts faced three primary criticisms, specifically, respresentativenes of the data, the adequacy of the IQ data, and the causal interpretation of the correlations. The current work addressed all these criticisms. First, this research explicitly addressed the departures from representativeness in the samples. The samples of public school children are not fully representative of all state children. An indicator (percent of White children in non-public schools) of the extent to which the state's sample is non-representative was offered. This indicator proved useful in accurately estimating the relationship between estimated state IQ and other important state variables. Thus, the explicit delineation of the departures from representativeness and the provision of an indicator of the degree of non-representativeness separates this research from that of Lynn and Vanhanen (2002). Second, the measure of estimated state IQ was the same in all states thus making estimated state IQ readily comparable across states. It is also a standardized measure of math and reading and is thus clearly a measure of cognitive ability. Thus, the measure of estimated state IQ may receive fewer criticisms than those measures offered by Lynn and Vanhanen (2002). Finally, this paper has been cautious in making causal inferences from the correlational data.

Appendix A. Sample items from the NAEP 4th grade math and reading exams

Sample NAEP 4th grade math item. Source: http://nces.ed.gov/nationsreportcard/itmrls/, accessed on August 4, 2006.



5. On the number line above, what number would be located at point *p*?

Answer:____________________

Sample NAEP 4th grade reading item. Source: http://nces.ed.gov/nationsreportcard/itmrls/, accessed on August 4, 2006. This item followed a reading passage title "How the Brazilian beetles got their coats."

8. The beetle chooses green and gold for the colors of her coat because they are

A) unusual colors
B) her favorite colors
C) the colors that rat wanted
D) the colors of her world

References

Agar, N. (2004). *Liberal eugenics: In defense of human enhancement.* Malden, MA: Blackwell.

Barnett, S. M., & Williams, W. (2004). National intelligence and the emperor's new clothes. *Contemporary Psychology*, *49*, 389−396.

Batty, G. D., & Deary, I. J. (2005). Education and mortality: A role for intelligence? *Journal of Epidemiology and Community Health*, *59*, 809−810.

Bouchard, T. J. (2004). Genetic influence on human psychological traits: A survey. *Current Directions in Psychological Science*, *13*, 148−151.

Boudette, N. E. (2006, March 10). Civics lesson. Kalamazoo, Mich., pegs revitalization on a tuition plan. *Wall Street Journal*.

Coleman, W. C., & Cureton, E. E. (1954). Intelligence and achievement: the 'jangle fallacy' again. *Educational and Psychological Measurement*, *14*, 347−351.

Cronbach, L. J. (1984). *Elements of psychological testing (4th ed.). New York: Harper & Row.*

Deary, I. J., Strand, S., Smith, P. & Fernandes C. (in press). Intelligence and educational achievement. *Intelligence*.

Ervik, A. O. (2003). IQ and the wealth of nations. *The Economic Journal*, *113*, 406−408.

Farrington, D. P. (2005). Childhood origins of antisocial behavior. *Clinical Psychology and Psychotherapy*, *12*, 177−190.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or *g*? The relationship between the SAT and general cognitive ability. *Psychological Science*, *15*(6), 373−378.

Gordon, R. A. (1987). SES versus IQ in the race-IQ-delinquency model. *International Journal of Sociology and Social Policy*, *7*, 30−96.

Gottfredson, L. S. (2004). Life, death, and intelligence. *Journal of Cognitive Education and Psychology*, *4*, 23−46.

Government Performance Project. (2006). Retrieved March 1, 2006 from http://results.gpponline.org/

Hart, C. L., Taylor, M. D., & Smith, G. D. (2005). Childhood IQ and all-cause mortality before and after age 65: Prospective observational study linking the Scottish Mental Survey 1932 and the Midspan studies. *British Journal of Health Psychology*, *10*, 153.

Hunt, E. (1995). *Will we be smart enough?* New York: Russell Sage Foundation.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

Klein, S. P., Hamilton, L. S., McCaffrey, D. F., & Stecher, B. M. (2000). What do test scores in Texas tell us? *Education Policy Analysis Archives*, *8*(49).

Koenig, K. A. (2006). *ACT and general cognitive ability.* Unpublished Master's Thesis. Department of Psychology. Case Western Reserve University.

Kristol, E. (1993). Picture perfect: The politics of prenatal testing. *First Things*, *32*, 17−24.

Lykken, D. T. (2001). Parental licensure. *American Psychologist*, *56*, 885−894.

Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations.* Westport, CT: Praeger.

McLaughlin, D., Gallagher, L., & Stancavage, F. (2004). Evaluation of bias correction methods for "Worst-case" selective non-participation in NAEP. *Commissioned by the NAEP Validity Studies (NVS) Panel.* Washington, DC: American Institutes for Research.

Murray, C. (2005). The inequality taboo. Commentary. article obtained from www.commentarymagazine.com/production/files/murray0905.html

National Center for Educational Statistics. (2006). Retrieved March 1, 2006 from http://www.nces.ed.gov

National Center for Health Statistics. (2005). *Health, United States, 2005, With Chartbook on Trends in the Health of Americans.* Hyattsville, Maryland: Author.

Richards, M. (2002). IQ and the wealth of nations. *Intelligence*, *30*, 574−576.

Roth, P. L., BeVier, C. A., Bobko, P., Switzer, F. S., III, & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, *54*, 297−330.

Scarr, S. (2000). Toward voluntary parenthood. *Journal of Personality*, *68*, 615−623.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262−274.

Strom, C. M. (2004, August, 12-17). Population-based carrier screening and prenatal diagnosis. *Medical Laboratory Observer.* Retrieved March 1, 2005 from www.mlo-online.com

Tzelgov, J., & Henik, A. (1991). Suppression situations in psychological research: Definitions, implications, and applications. *Psychological Bulletin*, *109*, 524−536.

United Health Foundation. (2006). Retrieved March 1, 2006 from http://www.unitedhealthfoundation.org. “America’s Health Rankings- 2005 Edition, © 2005 United Health Foundation. All Rights Reserved.”

U.S. Census Bureau (2006). Retrieved March 1, 2006 from http://factfinder.census.gov/home/saff/main.html?_lang=en

U.S. Center for Disease Control. (2006). Retrieved March 1, 2006 from http://www.cdc.gov

U.S. Department of Commerce Bureau of Economic Analysis. (2006). Retrieved March 1, 2006 from http://www.bea.gov

U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics. (2006). Retrieved March 1, 2006 from http://www.ojp.gov/bjs/dtdata.htm

Volken, T. (2003). IQ and the wealth of nations: A critique of Richard Linn and Tatu Vanhanen’s recent book. *European Sociological Review*, *19*, 411−412.

Wiener, J. (May 09, 2002). *Seeking diversity or academic status, colleges court Jews.* Jewish Telegraphic Agency, Retrieved March 1, 2006 from http://news.ncmonline.com/news/view_article.html?article_id=233