Methodological Notes[1]

Calculating the American HD Index

The American Human Development (HD) Index is calculated using a simple methodology that is replicated for each state, congressional district, metro area, and population group. First, a sub-index for each of the three components of the overall index—health, education, and income—is calculated, and each of the three components is weighted one-third in the index. This equal weighting is not arbitrary, but rather reflects a belief that these three basic building blocks of a life of freedom and opportunity are equally essential. Performance in each dimension is expressed as a value between 0 and 10 by applying the following general formula:

$$\text{Dimension Index} = \frac{\text{actual value} - \text{minimum value}}{\text{maximum value} - \text{minimum value}} \times 10$$

Goalposts for Calculating the American HD Index

For each of the three indices, goalposts are determined based on the range of the indicator observed for all possible groupings and also taking into account possible increases and decreases in years to come. In order to make the HD Index comparable over time, the health and education indicator goalposts do not change from year to year. The earnings goalposts are adjusted for inflation (please see the Income Index section below for more details). Because earnings data and the goalposts are presented in dollars of the same year, these goalposts reflect a constant amount of purchasing power regardless of the year, making Income Index results comparable over time.

INDICATOR	MAXIMUM VALUE	MINIMUM VALUE
Life expectancy at birth (years)	90	66
Educational attainment score	2.0	0.5
Combined gross enrollment ratio (%)	100	70
Median personal earnings (2009 dollars)*	$60,429	$14,283

* Earnings goalposts were originally set at $55,000 and $13,000 in 2005 dollars.

The American HD Index results from taking the simple average of the health, education, and income indices. Since all three components range from 0 to 10, the HD Index itself also varies from 0 to 10, with 10 representing the highest level of human development.

EXAMPLE:

Calculating the HD Index for the United States

HEALTH Index

Life expectancy at birth for the United States was 78.6 years in 2007. The Health Index is calculated as follows:

$$\text{Health Index} = \frac{78.6 - 66}{90 - 66} \times 10 = \mathbf{5.25}$$

EDUCATION Index

In 2008, 85 percent of U.S. adults had at least a high school diploma, 27.7 percent had at least a bachelor's degree, and 10.2 percent had a graduate degree. The Educational Attainment Score is 0.85 + 0.277 + 0.102 = 1.228. The Educational Attainment Index is then:

$$\text{Educational Attainment Index} = \frac{1.228 - 0.5}{2.0 - 0.5} \times 10 = \mathbf{4.86}$$

The combined gross enrollment ratio was 87.3 percent, and the Enrollment Index is then:

$$\text{Enrollment Index} = \frac{87.3 - 70}{100 - 70} \times 10 = \mathbf{5.76}$$

The Educational Attainment Index and the Enrollment Index are then combined to obtain the Education Index:

$$\text{Education Index} = \frac{2}{3}\mathbf{4.86} + \frac{1}{3}\mathbf{5.76} = \mathbf{5.15}$$

INCOME Index

Median earnings in 2008 were $29,755 (in 2009 dollars). The Income Index is then:

$$\text{Income Index} = \frac{\log(29{,}755) - \log(14{,}283)}{\log(60{,}429) - \log(14{,}283)} \times 10 = \mathbf{5.09}$$

HUMAN DEVELOPMENT Index

Once the dimension indices have been calculated, the HD Index is obtained by a simple average of the three indices:

$$\text{HD Index} = \frac{\mathbf{5.25} + \mathbf{5.15} + \mathbf{5.09}}{3} = \mathbf{5.17}$$

Geographic, Racial, and Ethnic Designations

Data in this book are presented for three geographic units: states, congressional districts, and Metropolitan Statistical Areas (MSAs).

Though Washington, DC, is not a state, for the purposes of this report it is treated as one. Doing so is common practice among other analyses of economic and social issues and follows the convention of the U.S. Census Bureau. Washington, DC, also has a larger population than Wyoming and is nearly as populous as Alaska, North Dakota, and Vermont.

The congressional districts in this book are those of the 111th Congress (2008–2010). Congressional districts are typically revised at the beginning of each decade based on the results of the decennial census. However, redistricting changes have occurred in both Georgia and Texas since 2005. Therefore, readers are advised not to compare congressional district data from *The Measure of America 2008–2009* to congressional district data in this volume for any of Georgia's thirteen congressional districts or for Texas Congressional Districts 15, 21, 23, 25, or 28.

MSAs are the designation for urban centers and their outlying areas as defined by the White House Office of Management and Budget. MSAs constitute counties grouped around an urban center of at least fifty thousand people plus outlying counties from which a substantial percentage of the population commute to the urban center. MSAs therefore include principal cities as well as outlying suburban and exurban areas.[2]

Racial and ethnic groups used in this analysis are based on definitions established by the White House Office of Management and Budget (OMB) and used by the Census Bureau, the Centers for Disease Control, and other government entities.[3] Since 1997 the OMB has recognized five racial groups and two ethnic categories. The racial groups include Native

Americans, Asian Americans, African Americans, Native Hawaiians and Other Pacific Islanders, and whites. The ethnic categories are Latino and not Latino. The Native American category includes Alaska Natives for the nation and in states where Alaska Natives reside. AHDP recognizes that Native Hawaiian and Other Pacific Islanders constitute one of the five racial groups recognized by the OMB. However, this group's very small population (about 428,000 in 2008 according to ACS estimates[4]) limits the availability of data for this group. We are therefore unable to provide a complete set of human development data for Native Hawaiians and Other Pacific Islanders at this time.

Health Index

The Health Index measures relative achievement in life expectancy at birth. Life expectancy at birth is calculated using data from two principal sources. Mortality data for 2007, the most recent year for which the data are available, were obtained by arrangement with the National Center for Health Statistics (NCHS) at the Centers for Disease Control and Prevention, and the National Association for Public Health Statistics and Information Systems Vital Statistics Cooperative Program. Bridged-race population estimates for the July 1, 2007, population (using Vintage 2008 data) were obtained from the CDC WONDER Database.

Life expectancy is calculated based on a widely used method developed by C. L. Chiang. This method involves the construction of abridged life tables that use population and mortality counts by age-group as inputs.

The Health Index is obtained by scaling the life expectancy at birth values using the maximum and minimum goalposts and is calculated as follows:

$$\text{Health Index}_i = \frac{LE_i - LE_{MIN}}{LE_{MAX} - LE_{MIN}} \times 10$$

where LE_i is the life expectancy at birth for unit *i* and LE_{MIN} and LE_{MAX} are the goalposts.

Estimation of Life Expectancy at Birth for Congressional Districts

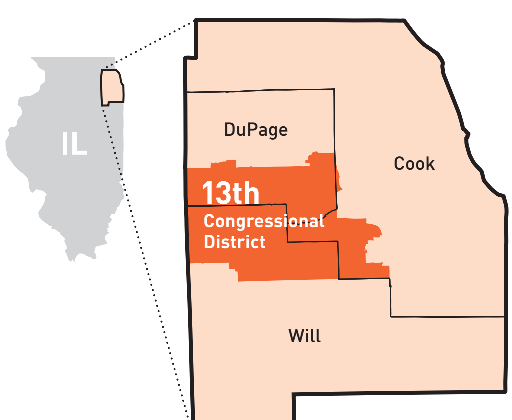
In the mortality data received from the Centers for Disease Control and Prevention, the state and county in which each decedent lived are flagged but the congressional district of residence is not. Therefore, life expectancy for congressional districts is estimated by apportioning death and population counts by county to congressional districts.

Congressional districts contain roughly 650,000 residents each and do not cross state lines. However, congressional districts do not necessarily conform to county lines within states. In order to determine how counties match up with congressional districts, a "geographic correspondence file" is generated to determine which counties fall within each congressional district and what percentage of the population of each county lives within that district.[5] Using this data, it is possible to allocate mortality and population counts at the county level to congressional districts.

In cases where mid-decade redistricting occurred, estimating congressional district life expectancy required an additional step. Changes that occurred in Georgia and Texas between 2005 and 2008, which affected all Georgia congressional districts and Congressional Districts 15, 21, 23, 25, and 28 in Texas, were not incorporated into the existing correspondence file, requiring AHDP to generate a new correspondence file for aligning counties to congressional districts in these two states. Using the MABLE/Geocorr application, a correspondence file was generated to match counties to Census tracts for Georgia and Texas based on 2008 population estimates. This was then compared to Geographic Relationship Tables for the 110th Congressional Districts in these two states in order to see how new congressional districts map onto existing Census tracts.[6] In the cases where Census tracts were split between congressional districts during this process, the populations of those tracts were split evenly between the two districts. Census tracts contain 8,000 people or less; thus, any distortions due to this arbitrarily even division of the population of split tracts are likely to be very small. As a test of accuracy, a comparison was made between the allocation factors generated by the MABLE/Geocorr application and those generated in-house by AHDP for the Texas districts unaffected by redistricting. These two sets of allocation factors were found to correlate very strongly (r = .955). This correlation suggests that life expectancy estimates made using county-to-congressional-district allocation factors generated by AHDP should be comparable to estimates made using allocation factors generated by the MABLE/Geocorr application.

The figure below illustrates this process. It shows the Thirteenth Congressional District in Illinois, comprising parts of Cook, DuPage, and Will counties. The proportion of each county's population that lives in the congressional district is computed, based on the Census block populations, and those proportions are then used to allocate death counts and population totals for the congressional district.

The Thirteenth Congressional District, Illinois



For this example, we have:

COUNTY	POPULATION	SHARE	DEATHS, < 1 YEAR	POPULATION, < 1 YEAR
Cook, IL	84,393	0.016	150	81,598
DuPage, IL	375,163	0.402	70	12,431
Will, IL	271,606	0.406	50	9,434

Population is the county's population residing in the congressional district; *share* is the percentage of the county's total population residing in the congressional district; *deaths* is the number of deaths of county residents in the age bracket (those are fictional numbers used for illustration purposes only, since the actual data are protected by a nondisclosure agreement); and *population* is each county's population in the age bracket. Thus, 1.6 percent of Cook County's residents, 40.2 percent of DuPage County's residents, and 40.6 percent of Will County's residents live in the Thirteenth Congressional District. The procedure then allocates 1.6 percent of the death counts in Cook County, 40.2 percent of the death counts in DuPage County, and 40.6 percent of the death counts in Will County to the target congressional district. The number of deaths in the < 1 year age bracket for the congressional district is given by

$$(0.016 \times 150) + (0.402 \times 70) + (0.406 \times 50) = 50.84$$

and the population in the same age bracket is given by

$$(0.016 \times 81{,}598) + (0.402 \times 12{,}431) + (0.406 \times 9{,}434) = 10{,}133$$

This procedure is repeated for all the age brackets, resulting in an abridged life table for the congressional district, which is then used to compute the life expectancy at birth.

In some instances, several congressional districts are entirely contained inside a single county; when this happens, the county's life expectancy at birth is assigned to all the congressional districts.

Life Expectancy Estimates for Asian Americans, Latinos, and Native Americans

One challenge in the calculation of life expectancy is the miscoding of race on death certificates, a surprisingly widespread occurrence affecting Asian Americans, Latinos, and Native Americans.[7] In order to calculate life expectancy for these groups, different methods are required to correct for errors in racial classification on death certificates.

The problem is particularly consequential for Native Americans. Drawing on studies undertaken by the National Center for Health Statistics and the Indian Health Service,[8] mortality counts have been adjusted using age group–specific correction factors based on current research about the prevalence of miscoding across the country and in specific states. Life expectancy estimates for this group are based on these adjusted mortality counts. Due to the small population size of Native Americans in the majority of states, and data inconsistencies in others, only twelve states could be included in this analysis.

For Asian Americans and Latinos, a ratio correction was applied in states in which these groups constitute a proportion of the state population that is less than half their share of the total national population. The cut-off values were thus as follows:

GROUP	% OF TOTAL POPULATION	CUT-OFF POINT BELOW WHICH CORRECTION WAS USED FOR STATES
Asian Americans	4.41% (13,413,976)	2.21%
Latinos	15.42% (46,891,456)	7.71%

Based on 2008 American Community Survey (ACS) 1-year population data, the correction was applied to mortality data for Asian Americans and Latinos in the states listed in the next column. As with state-level disaggregation in the other indices, life expectancy was not estimated for population groups smaller than 50,000 in any given state due to the statistical instability of working with survey-based estimates for small populations.

ASIAN AMERICANS	LATINOS
Florida*	Alabama
Indiana	Arkansas
Kansas	Delaware
Louisiana	Georgia*
Missouri	Indiana
North Carolina	Iowa
Ohio	Kentucky
Oklahoma	Louisiana
South Carolina	Maryland
Tennessee	Michigan
Utah	Minnesota
Wisconsin	Mississippi
	Missouri
	North Carolina
	Ohio
	Oklahoma
	Oregon*
	Pennsylvania
	Rhode Island*
	South Carolina
	Tennessee
	Virginia
	Wisconsin

* Although populations in these states are technically above the cut-off point for the application of the correction, the correction was applied nonetheless in order to ensure a plausible result.

To apply the correction method to data for Asian Americans or Latinos in a given state, for each age group, the ratio of the death rate for that group at the national level to the corresponding rate for the total population (i.e., all racial/ethnic groups combined) is calculated. Age-specific death rates for the total population in that state are then multiplied by this national ratio. In this way, for each age group, the ratio of the death rate for the state's race/ethnic group at issue to the rate for the state's total population is identical to the corresponding ratio in the national population. An identical method is used to estimate life expectancy at birth for racial and ethnic groups within certain MSAs in which Asian Americans and Latinos constitute a small percentage of the total population. This method is employed for Asian Americans in the Miami–Fort Lauderdale–Pompano Beach MSA, and for Latinos in the Atlanta–Sandy Springs–Marietta, Boston-Cambridge-Quincy, and Washington-Arlington-Alexandria MSAs. These estimates are one way to approximate life expectancies for racial and ethnic groups for which existing data are too flawed to permit actual calculations. Readers are advised to bear this in mind when comparing these estimates with life expectancy calculations for other groups.

Education Index

The Education Index is based on two sub-indices: an Educational Attainment Index and an Enrollment Index. The **Educational Attainment Index** measures the overall level of educational attainment achieved by the adult population. It takes into account the percentage of the population age 25 years and older who have earned at least a high school diploma or equivalent, at least a bachelor's degree, or an advanced degree (master's, professional, doctoral, etc.). Each category represents the percentage of the adult population who have achieved at least that level of attainment, meaning that the percentage of the population 25 and over with a master's degree necessarily includes those with a bachelor's degree and a high school diploma or its equivalent. To calculate the Educational Attainment Index, first an Attainment Sum is determined by adding the percentage of the population 25 and older with at least a high school diploma or equivalent, the percentage with at least a bachelor's degree, and the percentage with an advanced degree. Those who have earned an associate degree or those who have completed some college without earning a degree are counted in the "at least high school" category. The Educational Attainment Index is calculated as follows:

$$\text{Educational Attainment Index}_i = \frac{EAS_i - EAS_{MIN}}{EAS_{MAX} - EAS_{MIN}} \times 10$$

where EAS_i is the Educational Attainment Score for unit i and EAS_{MIN} and EAS_{MAX} are the goalposts.

The **Enrollment Index** is based on a gross enrollment calculation that takes into account the total number of students enrolled in school (of any age at any level) divided by the total school-aged population of 3- to 24-year-olds (inclusive). Therefore,

$$\text{Gross Enrollment Ratio}_i = \frac{ENR_i}{P3TO24_i}$$

where ENR_i is the population of any age enrolled in school at any level and $P3TO24_i$ is the population between the ages of 3 and 24. The Enrollment Index is then calculated:

$$\text{Enrollment Index}_i = \frac{GER_i - GER_{MIN}}{GER_{MAX} - GER_{MIN}} \times 10$$

where GER_i is the Educational Attainment Score for unit i and GER_{MIN} and GER_{MAX} are the goalposts. If the Gross Enrollment Ratios exceed 100 percent, as can happen when large numbers of older students are enrolled in school, the Gross Enrollment Ratio is capped at 100 percent for the purposes of calculating the Enrollment Index.

Finally, these two components are combined into the Education Index. In order to reflect the relative ease of enrolling students in school compared to the completion of a meaningful course of education (signified by the attainment of degrees), a two-thirds weight is applied to the Attainment Index and a one-third weight to the Enrollment Index to calculate the final Education Index as follows:

$$\text{Education Index}_i = \frac{2}{3} EAI_i + \frac{1}{3} EI_i$$

where EAI_i is Educational Attainment Index, and EI_i is Enrollment Index.

Attainment data for the United States are obtained from the ACS using attainment by education level and population from form B15002 (Sex by Educational Attainment for Population 25 Years and Over), B15002B, B15002C, B15002D, B15002H, and B15002I (same, for Black or African American Alone, Native American and Alaska Native Alone, Asian

Alone, White Non-Hispanic, and Hispanic).

Enrollment data for the United States as a whole, for individual states, and for states by race or by gender (but not by race and gender) were obtained from ACS tables B14001 (School Enrollment by Level of School for the Population 3 Years and Over), B14001B, B14001C, B14001D, B14001H, and B14001I (same, for Black or African American alone, Native American and Alaska Native Alone, Asian Alone, White Non-Hispanic Alone, and Hispanic). Enrollment data for racial and ethnic groups broken down by gender were obtained by AHDP analysis of the ACS 2008 1-year Public Use Microdata Sample (PUMS), queried using the U.S. Census Bureau DataFerrett tool.

Income Index

The Income Index is calculated as follows:

$$\text{Income Index}_i = \frac{\log(y_i) - \log(y_{MIN})}{\log(y_{MAX}) - \log(y_{MIN})} \times 10$$

where y_i is the Median Earnings for unit i and y_{MIN} and y_{MAX} are the goalposts.

Median personal earnings data for the United States were obtained from ACS tables B20017 (Median Earnings by Sex by Work Experience for the Population 16+ Years with Earnings), B20017B, B20017C, B20017D, B20017H, and B20017I (same table for Black and African American Alone, Native American and Alaska Native Alone, Asian Alone, White Non-Hispanic Alone, and Hispanic). Median personal earnings reflect the median of the sum of wages, salaries, and net income from self-employment before deductions for taxes, and social contributions for the population age 16 and over with earnings.

Inflation adjustments. Comparing earnings from different years requires an adjustment to account for the depreciation of the purchasing power of any dollar amount due to inflation. The Consumer Price Index (CPI) as calculated by the Bureau of Labor Statistics (BLS) was used to convert dollars of different years to 2009 dollars for the purposes of this report. Following the recommendation of the U.S. Census Bureau,[9] the CPI research series using current methods (CPI-U-RS) was used to construct conversion factors for converting dollars of one year to another.

Error Margins

All of the data used to calculate the American HD Index besides life expectancy comes from the American Community Survey (ACS), an annual survey conducted by the U.S. Census Bureau that samples a small, randomly selected percentage of the population. Although the ACS is an excellent resource, as with any survey, there is some degree of sampling error. Thus, not all differences between two places or groups reflect the true difference between those places or groups. Thus, comparisons between similar values on any indicator, especially for small populations, should be made with caution since these differences may not always be statistically significant. Standard error and margin of error data for American HD Index values can be found at http://www.measureofamerica.org/report2010-11methods/. Readers interested in testing the statistical significance of estimates presented in this report are encouraged to view the supplemental web-content on error margins.

Difference between the American HD Index and the UNDP HD Index

The original HD Index was created by the United Nations Development Programme (UNDP) and is published annually in the *Human Development Report*. This composite index was created to measure human development in all countries of the world, ranging from very-low-income countries in sub-Saharan Africa to high-income countries in Europe, North America, and elsewhere. Thus, some of the indicators used are not well suited to measuring human development in an advanced industrialized economy like the United States. Nor are the goalposts for the

UNDP HD Index helpful for assessing U.S. well-being as they have been set to accommodate a very wide range of conditions—for instance, literacy rates ranging from less than 30 percent in Chad, Mali, and Burkina Faso to close to 100 percent in many other countries.

The American Human Development Project modified the UNDP HD Index to create the American HD Index. The American HD Index follows the same principles as the UNDP HD Index, and measures the same three basic dimensions of human development—a long and healthy life, access to knowledge, and a decent standard of living—but it has been adapted in order to better reflect the U.S. context.

The table below lists the indicators used in the American HD Index and the UNDP HD Index:

DIMENSION	INDICATOR AMERICAN HD INDEX	INDICATOR UNDP HD INDEX
Long and Healthy Life	Life expectancy at birth	Life expectancy at birth
Access to Knowledge	Educational attainment Gross enrollment ratio	Adult literacy rate Gross enrollment ratio
Decent Standard of Living	Median personal earnings	GDP per capita

In the health dimension, the same indicator is used (life expectancy at birth), but the goalposts are changed. The UNDP HD Index uses goalposts of 25 years (minimum) and 85 years (maximum) to accommodate the enormous gap in life expectancy found in countries around the world. For the American HD Index, the goalposts were set at 66 years and 90 years, a range that accommodates the variations across all groupings considered in *The Measure of America*. Since life span in the United States is nowhere near the lower limit of 25 years set in the standard HD Index, using the standard HD Index goalposts would cluster all Health Index values around the maximum value, providing very little differentiation among states, congressional districts, and so on.

In the knowledge dimension, adult literacy rate was replaced with an educational attainment index. Adult literacy is a relevant indicator in a global context, where low-income countries still have very high illiteracy rates, but is largely irrelevant for developed nations, where most of the adult population has basic reading and writing skills and the labor market demands increasingly sophisticated skills. Functional literacy (the ability to read, write, and speak in English, and compute and solve problems at levels of proficiency necessary to function on the job and in society, achieve one's goals, and develop one's knowledge and potential) would be a good indicator, but suffers from severe data availability problems. Thus, the educational attainment index was used. It captures the overall educational level of the population and is a good indicator of the extent to which people have the skills necessary to carry out the essential tasks of daily life in the United States, lead lives of choice and value, and meet the increasing demands of the labor market.

The other knowledge indicator, school enrollment, which is the combined gross enrollment ratio, is the same in both the American HD Index and the UNDP HD Index, with a slight modification. The enrollment ratio in the American HD Index includes nursery school and pre-kindergarten, and the age group used in the denominator of the enrollment ratio has been adjusted to accommodate this (the range begins with age 3). The goalposts were also changed, from 0 to 100 percent in the UNDP HD Index to 70 to 100 percent in the American HD Index, to reflect the ranges observed in all American HD Index groupings.

In the standard of living dimension, GDP per capita was replaced by median personal earnings. For relatively closed economies, such as those of countries, GDP per capita is a good indicator of the income appropriated by the local population. However, in smaller geographical areas within a country, such as states and congressional districts, economies are much more open, and substantial portions of the income generated within a community are used to remunerate production factors owned by persons who do not reside in that community (e.g., profits from a large manufacturing plant located in the community). Using such an indicator for states and congressional districts therefore would not adequately represent the income available to local residents. Also, there is no way to measure such an indicator for racial and ethnic groups or for men or women alone. Using median personal earnings (rather than household earnings) allows for the assessment of the difference in command over economic resources between women and men; median earnings also allow for disaggregation by race and ethnicity as well as by congressional district and state.

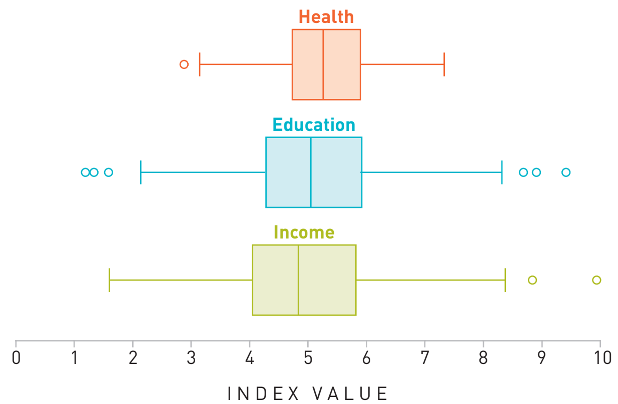
As a result of these modifications, the American HD Index and the UNDP HD Index are not comparable. In order to prevent any comparison attempts, the American HD Index varies from 0 to 10, whereas the UNDP HD Index varies from 0 to 1.

Balancing the American HD Index Components

In any composite index, each component of the index should contribute equally to the overall index value. The American HD Index assigns equal weights to each of the three components of the index, but how much each of these components affects the final score cannot be assumed to be equal based on this alone. If the distributions of the scores in the three components are not similar, then some components may end up having an **implicit weight** that results in one component having more of an influence on final scores than other components. How equally balanced the components of the HD Index are can be assessed by looking at the distribution of scores using box plots and also by looking at descriptive statistics and regression coefficients for these components.

A look at the distribution of scores for congressional districts on the three component indices of the American HD Index (displayed below) shows that these components are relatively well-balanced. Scores cluster around 5, and the ranges and distributions are similar. There are a few outliers, however, and the range of scores on the Health Index is narrower than the ranges of scores on the other two indices; the goalposts for the Health Index had to be set wide in order to accommodate large variations in life expectancy among racial and ethnic groups, and between women and men. As variations in life expectancy among congressional districts and states were not quite as extreme, scores on this index tend to clump closer to the midpoint.

Box Plots for the Component Indices of the American HD Index (Congressional Districts)



A further test of how equally each component index contributes to the final HD Index Value is to look at both the average values for each component index and also to run a simple linear regression with the final HD Index as the dependent variable and the three components as the independent variables. If the unstandardized coefficients in the resulting regression analysis are all roughly equal, each component is playing an equal role in contributing to the final HD Index score.

American HD Index Components

COMPONENT	AVERAGE COMPONENT INDEX VALUE	AVERAGE CONTRIBUTION
By State		
Health	5.15	33.4
Education	5.13	33.3
Income	4.84	33.3
By Congressional District		
Health	5.27	35.0
Education	5.15	33.0
Income	4.99	32.6

The tables above suggest that the average values and average contributions to the final HD Index of each of the component indices are roughly equal. Therefore, the HD Index is relatively well-balanced; none of the three component indices is having a disproportionately large impact on the final HD Index score.