Methodological Notes

The American Human Development Index

The modified American Human Development Index measures the same three basic dimensions as the standard HD Index, but it uses different data in order to better reflect the U.S. context and to maximize available data. All data come from official U.S. government sources. The most recent year for which data are available is 2005, owing to the typical lag time of two to three years.

In the American Human Development Index:

- **A long and healthy life** is measured using life expectancy at birth, calculated from mortality data from the Centers for Disease Control and Prevention, National Center for Health Statistics, and population data from the U.S. Census Bureau, 2005.
- **Access to knowledge** is measured using two indicators: school enrollment for the population age three and older, and educational degree attainment for the population twenty-five years and older. Both indicators are from the American Community Survey, U.S. Census Bureau, 2005.
- **Decent standard of living** is measured using median earnings from the American Community Survey, U.S. Census Bureau, 2005.

Before the HD Index itself is calculated, an index needs to be created for each of these dimensions. To calculate these indices—the health, education, and income indices—minimum and maximum values (goalposts) are chosen for each underlying indicator. Performance in each dimension is expressed as a value between 0 and 10 by applying the following general formula:

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

The HD Index is then calculated as a simple average of the dimension indices.

Health Index

The Health Index measures the relative achievement in life expectancy at birth. Life expectancy is calculated by constructing a life table. Life tables are a series of columns of data; the only raw data needed are the population and the number of deaths. All other columns of data and the expectation of life are calculated from these.

Life tables in this report have been constructed using Chiang's method of abridged life tables. Abridged life tables aggregate deaths and population data into age groups, rather than using single year of age as do complete life tables. The age groups used were under 1, 1–4, 5–9, . . . , 80–84, and 85 and over. The Chiang method is well established and widely used internationally.[196]

Death data were obtained from the National Center for Health Statistics (NCHS). NCHS data record all deaths occurring in the United States, and include cause of death, county of residence, race, sex, and age. The public use mortality files made available by NCHS do not include county identifiers for deaths in counties with fewer than one hundred thousand people, for confidentiality reasons. We obtained county identifiers for all deaths through a special request to NCHS.[197]

Population data are the bridged-race population estimates of the July 1, 2005, population, produced by the U.S. Census Bureau in collaboration with NCHS.[198]

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

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

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

The observed ranges for life expectancy at birth were:

Grouping	Minimum	Maximum
States	73.8 (District of Columbia)	81.7 (Hawaii)
Race/Ethnicity	73.0 (African Americans)	86.3 (Asians)
Race/Ethnicity × Gender	69.4 (African American males)	88.8 (Asian females)
Congressional Districts	72.6 (5th CD, Kentucky)	82.9 (8th CD, Virginia)

The goalposts are determined based on the range of the indicator observed on all possible groupings and also taking into account possible increases and decreases in years to come, and adjusted in order to achieve a balance in the final index (see **Balancing the American HD Index Components**). The goalposts were set at 66 (minimum) and 90 (maximum).

Education Index

The Education Index measures the relative achievement in both educational attainment and combined primary, secondary, and tertiary gross enrollment. An index for educational attainment and one for gross enrollment are calculated, and the two indices are combined to create the Education Index, with two-thirds weight given to educational attainment and one-third weight to gross enrollment.

EDUCATIONAL ATTAINMENT INDEX

The Educational Attainment Index measures the overall educational level of the adult population, and is based on a very simple premise—that more education is better. In addition to the obvious benefits of education for greater earnings potential for individuals, higher educational attainment is associated with many other benefits as well. These can include better health, increased civic participation, increased ability to adjust to change, clearer self-identity, and greater social capital passed on to their children. For society, a better-educated population is associated with reduced dependency on public support programs, lower crime rates, political stability, and environmental benefits.

The Educational Attainment Index utilizes three indicators: percentage of the population twenty-five years and older who have completed high school (a high school diploma or equivalent, such as GED), percentage of the population twenty-five years and older with a bachelor's degree (does not include community college and associate degree) and percentage of the population twenty-five years and older with a graduate degree (master's, professional, or doctoral degree). Each category represents the percentage of the population that has attained *at least* that educational level. Thus, the percentage of high school graduates includes those with a bachelor's degree or higher; and the percentage of the population with a bachelor's degree includes those with a graduate degree.

An Educational Attainment Score is computed by adding the three indicators. This way, those with a graduate degree are counted three times, because they also must have a bachelor's degree and a high school diploma; and those with a bachelor's degree are counted twice, because they must have a high school diploma. Thus, if two communities have the same percentage of high school graduates, but one has a higher percentage of persons with a bachelor's degree, the one with more bachelor's degrees will get a higher Educational Attainment Score. The minimum value for the Educational Attainment Score is 0 —100 percent of the adult population with less than a high school diploma—and the maximum value is 3 —100 percent with a graduate degree.

The data source is the American Community Survey, tables B15002 (Sex by Educational Attainment for the Population 25 Years and Over), B15002B, B15002C, B15002D, B15002H, and B15002I (same, for Black or African American Alone, American Indian and Alaska Native Alone, Asian Alone, White Non-Hispanic Alone, and Hispanic).

The Educational Attainment Index is obtained by scaling the Educational Attainment Score values using the minimum and maximum goalposts:

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

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

The observed ranges for the Educational Attainment Score were:

Grouping	Minimum	Maximum
States	1.036 (Mississippi)	1.540 (District of Columbia)
Race/Ethnicity	0.755 (Latinos)	1.546 (Asians)
Race/Ethnicity × Gender	0.739 (Latino males)	1.656 (Asian males)
Congressional Districts	0.608 (20th CD, California)	1.804 (14th CD, New York)

The goalposts were set at 0.5 (minimum) and 2.0 (maximum).

ENROLLMENT INDEX

The Enrollment Index measures the relative achievement in combined primary, secondary, and tertiary gross enrollment. The indicator used is the combined gross enrollment ratio for primary, secondary, and tertiary levels. This is computed as the ratio of the number of students—of any age—enrolled at all three levels to the size of the population of the official age group corresponding to these levels. For the United States, the appropriate age group would be approximately five to twenty-four years of age. However, both the ACS and the decennial Censuses, when presenting enrollment data by race/ethnicity, combine nursery school, prekindergarten, and kindergarten together. Thus, the enrollment ratio has to take nursery school and prekindergarten into account, and the age group used to calculate the ratio is set as three to twenty-four years of age.

The enrollment ratio is a *flow* variable; it gives an indication of the future level of educational attainment for a given community, as opposed to the educational attainment indicator, which measures the present *stock* of education. Even though school attendance is mandatory in the primary and secondary levels, there is considerable variation in the enrollment ratio, due to high school dropout rates and enrollment in preschool and tertiary levels.

The data source is the American Community Survey, 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, American Indian and Alaska Native Alone, Asian Alone, White Non-Hispanic Alone, and Hispanic); table B14002 (Sex by School Enrollment by Type of School by Age for the Population 3 Years and Over); tables B01001 (Sex by Age), B01001B, B01001C, B01001D, B01001H, and B01001I (same, for Black or African American Alone, American Indian and Alaska Native Alone, Asian Alone, White Non-Hispanic Alone, and Hispanic).

The Gross Enrollment Ratio is given by:

$$\text{Gross Enrollment Ratio}_i = \frac{ENR_i}{P3T024_i}$$

where ENR_i is the population (of any age) enrolled in school, at all levels, and $P3T024_i$ is the population from three to twenty-four years of age. The Enrollment Index is then obtained by scaling the gross enrollment ratio using the minimum and maximum goalposts:

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

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

The observed ranges for the Gross Enrollment Ratios in the 2005 ACS were:

Grouping	Minimum	Maximum
States	80.2% (Idaho)	99.8% (District of Columbia)
Race/Ethnicity	78.8% (Latinos)	102.3% (Asians)
Race/Ethnicity × Gender	76.1% (Latino males)	106.0% (Asian females)
Congressional Districts	70.5% (4th CD, Arizona)	111.6% (8th CD, California)

The goalposts were set at 70 percent (minimum) and 100 percent (maximum). Since these are gross enrollment ratios, and use the population of *any age* enrolled in school, it is possible to obtain ratios greater than 100 percent, due to over-age enrollment. When this happens, the value is capped at 100 percent, so the Enrollment Index is never greater than 10.

The Educational Attainment and the Enrollment Indices are combined to form the Education Index, with two-thirds weight to Educational Attainment and one-third to Enrollment:

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

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

Income Index

In the HD Index, income serves as a surrogate for all the dimensions of human development not reflected in a long and healthy life and access to knowledge. Following the standard UNDP methodology, a logarithmic transformation is applied to the income indicator, to reflect a decreasing returns to scale assumption—an income increase at lower income levels has a greater impact on the overall level of material well-being than the same increase at higher income levels. Once a high enough income threshold is reached, additional income increases will have very little impact on material well-being.

Given the tremendous gaps in wealth and assets in the United States, which are much larger than the income gaps, especially across racial lines, a wealth indicator would be a worthy addition to the income index. Wealth is one of the key drivers of the intergenerational transmission of advantage, and an asset cushion is often what separates those who can remain middle class from those who experience periods of poverty in the case of a shock such as short-term unemployment or a serious illness. Unfortunately, wealth data are scarce and not available at the subnational level, and had to be left out.

The ACS offers several income measures; median earnings of the population sixteen years and older was chosen because it is the only income measure that is available for all the groupings used in this report (gender, race/ethnicity, gender by race/ethnicity, regions, states, congressional districts). Earnings are defined as the sum of wage or salary income and net income from self-employment, and do not include interest, dividends, rental income, Social Security income, and public assistance income. Only individuals with earnings are included in the computation of the median; "zero earners" are excluded.

The data source is the American Community Survey, tables B20017 (Median Earnings by Sex by Work Experience for the Population 16+ Yrs with Earnings), B20017B, B20017C, B20017D, B20017H,

and B20017I (same, for Black or African American Alone, American Indian and Alaska Native Alone, Asian Alone, White Non-Hispanic Alone, and Hispanic).

The Income Index is obtained by scaling the median earnings values using the minimum and maximum goalposts:

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

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

The observed ranges in the 2005 ACS were:

Grouping	Minimum	Maximum
States	$21,472 (Montana)	$36,948 (District of Columbia)
Race/Ethnicity	$20,255 (Latinos)	$31,518 (Asians)
Race/Ethnicity × Gender	$16,147 (Latina Females)	$37,269 (White Males)
Congressional Districts	$16,767 (20th CD, California)	$51,139 (14th CD, New York)

The goalposts were set at $13,000 (minimum) and $55,000 (maximum).

Calculating the Human Development Index

The HD Index is obtained by the simple average of the health, education, and income indices:

$$\text{HD Index}_i = \frac{\text{Health Index}_i + \text{Education Index}_i + \text{Income Index}_i}{3}$$

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

1. HEALTH Index

The life expectancy at birth for the U.S. was 78.0 years in 2005. The Health Index is given by

$$\text{Health Index} = \frac{78 - 66}{90 - 66} \times 10 = \mathbf{5.00}$$

2. EDUCATION Index

In 2005, 84.2 percent of U.S. residents had at least a high school diploma, 27.2 percent had at least a bachelor's degree, and 10.0 percent had a graduate degree. Then, the Educational Attainment Score is 0.842 + 0.272 + 0.100 = 1.214. The Educational Attainment Index is then

$$\text{Educational Attainment Index} = \frac{1.214 - 0.5}{2.0 - 0.5} \times 10 = \mathbf{4.76}$$

The combined gross enrollment ratio was 86.8 percent, and the Enrollment Index is then

$$\text{Enrollment Index} = \frac{86.8 - 70}{100 - 70} \times 10 = \mathbf{5.61}$$

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

$$\text{Education Index} = \frac{2}{3}\mathbf{4.76} + \frac{1}{3}\mathbf{5.61} = \mathbf{5.04}$$

3. INCOME Index

Median earnings in 2005 were $27,299. The Income Index is then

$$\text{Income Index} = \frac{\log(27{,}299) - \log(13{,}000)}{\log(55{,}000) - \log(13{,}000)} \times 10 = \mathbf{5.14}$$

4. HUMAN DEVELOPMENT Index

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

$$\text{HD Index} = \frac{\mathbf{5.00} + \mathbf{5.04} + \mathbf{5.14}}{3} = \mathbf{5.06}$$

Differences between the American HD Index and the Standard HD Index

The standard HD Index, created by UNDP and published in the annual Human Development Reports, was developed to measure human development in all countries of the world, ranging from very-low-income countries in sub-Saharan Africa to high-income OECD countries. Thus, some of the indicators used (and the goalposts for each indicator) are not very well suited to measure human development in a high-income country like the United States, since they have to accommodate a very wide range.

The American HD Index follows the same principles of the standard HD Index, and measures the same three basic dimensions of human development—health, access to knowledge, and standard of living—but has been modified in order to better reflect the context of a developed country.

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

	Indicator	
Dimension	**AMERICAN HD Index**	**STANDARD HD Index**
Health	Life expectancy at birth	Life expectancy at birth
Knowledge	Educational attainment Gross enrollment ratio	Adult literacy rate Gross enrollment ratio
Standard of living	Median earnings	GDP per capita

In the health dimension, the same indicator is used (life expectancy at birth), but the goalposts are changed. The standard HD Index uses goalposts of twenty-five years (minimum) and eighty-five years (maximum), to accommodate the enormous gap in life expectancy around the world. For the American HD Index, the goalposts were set at sixty-six years and ninety years, a range that accommodates the variations across all groupings used in the Report. Since life expectancies in the United States are higher than in most countries, and do not go anywhere near the lower limit of twenty-five years set in the standard HD Index, using the standard HD Index goalposts would cluster all Health Index values around the maximum value of 10, providing very little differentiation among states, congressional districts, and so on.

In the knowledge dimension, adult literacy rate was replaced by the educational attainment index. Adult literacy is a relevant indicator in a global context, where low-income countries still have very high illiteracy levels, 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 how well any given population is prepared to satisfy an increasingly demanding labor market.

The other knowledge indicator, combined gross enrollment ratio, is the same in both the American HD Index and the standard HD Index, but with a slight modification—the enrollment in the American HD Index includes nursery school and prekindergarten, and the age group used in the denominator of the enrollment ratio has been adjusted to accommodate this. The goalposts were also changed, from 0 to 100 percent in the standard HD Index to 70 to 100 percent in the American HD Index, in order to reflect the ranges observed on all American HD Index groupings.

In the standard of living dimension, GDP per capita was replaced by median earnings. For relatively closed economies, such as countries, GDP per capita is a good indicator of the income appropriated by the local population. However, for smaller geographical areas within a country, such as states and congressional districts, which are much more open economies, substantial portions of the

income generated within the 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), and thus do not adequately represent the income available to local residents.

As a result of these modifications, *the American HD Index and the standard HD Index are not comparable.* In order to reduce comparisons, the American HD Index varies from 0 to 10 instead of from 0 to 1, as does the standard HD Index.

Balancing the American HD Index Components

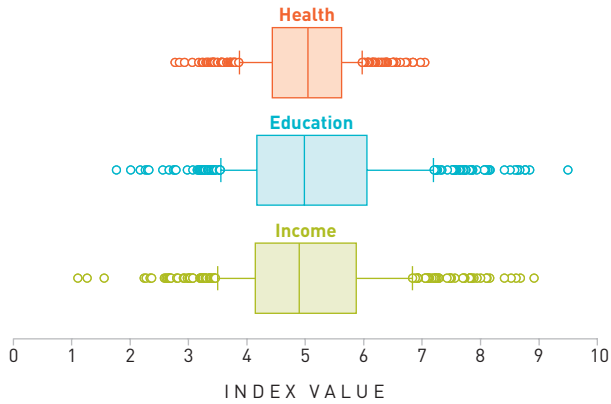
Ideally, the components of a composite index should be balanced; that is, on average, each component should contribute equally to the final index (assuming an equal weights system). The HD Index uses equal weights for the components, but this is not enough to warrant a balanced index. If the distributions of the components are not similar, some of the components may end up contributing to the final index more than others—in effect, making the *implicit weights* not equal. For instance, if one of the components has a range of 7 to 9, with the values clustered around 8, and the others have an average value of 5, the component with higher values will have a higher implicit weight, even though all components should have equal weights.

One way to ensure that no single component has a disproportionate implicit weight is to examine the distributions of the components and adjust the goalposts, so that the ranges and the distributions are similar.

The goalposts for each component of the American HD Index were carefully adjusted in order to accommodate the ranges of the observed values, allow for growth in future years, and create a balanced final index.

The resulting indices have median values close to 5, and well-balanced ranges and distributions, as shown by the box-and-whisker plots (the plots are for the congressional district indices; the state indices show a very similar pattern).

Box and Whisker Plot



The Health Index has a slightly more compressed range, because the goalposts have to accommodate the values for life expectancy by gender and race/ethnicity, which have a wider range than the ranges observed for states and congressional districts.

The average values and average contributions of each component, both for states and congressional districts, show a well-balanced index, with each component contributing equally, on average, to the final index.

American HD Index Components

by STATE

Component	Average Value	Average Contribution
Health	4.92	33.1%
Education	4.97	33.4%
Income	5.01	33.5%

by CONGRESSIONAL DISTRICT

Component	Average Value	Average Contribution
Health	5.02	33.5%
Education	5.03	32.7%
Income	5.20	33.8%

Estimation of Life Expectancy at Birth for Congressional Districts

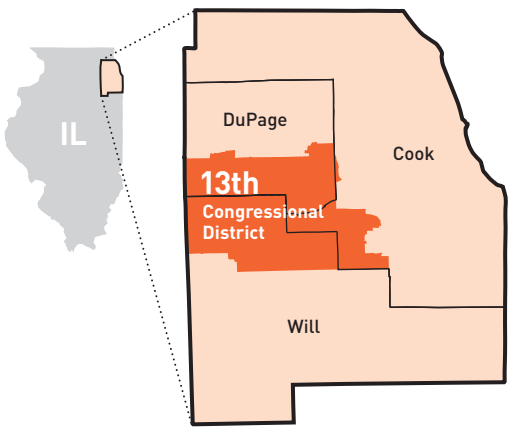
The Multiple Cause of Death data used in the calculation of life expectancy at birth does not contain congressional district identifiers, so it is not possible

to compute life expectancies directly for the congressional districts; they have to be estimated.

A procedure was created to build abridged life tables for each congressional district based on county level data. Congressional district boundaries do not conform to county boundaries; some congressional districts lie entirely within a single county, while others comprise parts of several different counties. The procedure generates a "geographic correspondence file" between congressional districts and counties, indicating what proportion of the congressional district's population lives in each county. This is done using Census blocks, the smallest geographic unit utilized by the Census Bureau. Then the death counts and population totals for each county are allocated to congressional districts, based on the allocation factors from the correspondence file. This creates the life tables for each congressional district, which allow the estimation of the life expectancies.

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 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.

The Geographic Correspondence File was generated by the MABLE/Geocorr application, developed by John Blodgett, from the University of Missouri St. Louis, and jointly owned by Blodgett and CIESIN (Consortium for International Earth Science Information Network, at Columbia University), whom we gratefully acknowledge.