UNEVEN OPPORTUNITIES: TECHNICAL APPENDIX

The Metropolitan Washington Healthy Places Index (HPI) Methods and Data

Derek Chapman, PhD[a]

David Wheeler, PhD[b]

Latoya Hill, MPH[a]

Steven Woolf, MD, MPH[a]

[1]Center on Society and Health, Virginia Commonwealth University

[2]Department of Biostatistics, Virginia Commonwealth University

VCU Center on Society and Health

October 2018

WHAT IS THE HEALTHY PLACES INDEX?

For many years, the counties in metropolitan Washington, D.C. have ranked among the healthiest in Maryland and Virginia.[1] But the health status of the local population is not uniform across the region. In fact, the statistics of individual neighborhoods vary dramatically. In February 2017, the Center on Society and Health (CSH) at Virginia Commonwealth University (VCU) was commissioned by the Metropolitan Washington Council of Governments (COG) Health Officials Committee (HOC) to examine life expectancy across the region's 1,223 census tracts and compute an index that quantifies the relative impact of various place-based determinants on life expectancy.

The project was part of a larger initiative of *Region Forward*[2], the COG strategic plan for the region. To support this effort, the HOC sought to gather statistics on two health metrics—life expectancy and quality of life—and consulted with CSH, which had longstanding experience with calculating life expectancy at the census tract level. The CSH team explained the methodological challenges and lack of data sources for measuring quality of life by census tract, but instead recommended an assessment of the social determinants of health, many of which aligned with *Region Forward* domains (e.g., housing, transportation). It proposed developing a composite measure (the Healthy Places Index [HPI]) that predicted life expectancy and could be used to visualize, through mapping, how the opportunity for good health varies across the census tracts of the metropolitan Washington region.

This document summarizes the methods used to calculate life expectancy at the census tract level, the social determinants of health that were collected for each census tract, the methods used for calculating the HPI, and the geographical footprint used for the analysis.

LIFE EXPECTANCY

Life expectancy at birth was calculated not only to provide the first of the two metrics of interest to the HOC but also as the dependent variable for validating the second metric, the HPI. Life expectancy at birth is the average number of years that a newborn can expect to live in a given area based on prior mortality patterns observed in that area. Life expectancy at birth for each census tract in the region was determined using mortality data that were previously geocoded by the health department responsible for recording deaths for that jurisdiction (i.e., District of Columbia, Maryland, or Virginia). Life expectancy data for census tracts in northern Virginia were provided by the Virginia Department of Health (based on 2007-2013 death data). Tract-level life expectancy for Maryland and Washington D.C. areas were computed by VCU using the most recently available 10 years of death data (Maryland: 2005-2014, District of Columbia: 2006-2015), which were provided by the Maryland Department of Health and the District of Columbia Department of Health, respectively.

Population counts for Virginia were obtained from the 2010 census, since the death data were centered around 2010. Decennial 2010 population counts were also used for Maryland and the District of Columbia since their midpoints (District of Columbia = 2010.5, Maryland = 2009.5) were almost exactly 2010. We preferred the decennial census counts over the American Community Survey data, which would have introduced much larger standard errors in the population estimates. For all areas, deaths were aggregated into 19 five-year age groups (0, 1-4, 5-9, 10-14, …80-84, 85+ years) by the decedent's residential census tract. The average number of deaths across the death years was computed in order to match the single year of population data used (2010).

All life expectancy calculations were completed using the adjusted Chiang II abridged life table method.[3] Death and population counts for age groups with zero deaths were replaced with the corresponding death and population counts for the county (or independent city in Virginia) that contained the tract with the zero count. Population data and death counts for census tracts that split or combined in 2010 were allocated based on the population distribution. Census tracts with ten or more missing age categories, and those with greater than 40% of the population living in group quarters in tracts with populations of less than 5,000, were excluded from the analysis.

Despite aggregating multiple years of death data, death and population counts in many census tracts remained small, resulting in large standard errors for life expectancy. In the District of Columbia and Maryland, the standard errors were typically between 2.0 and 3.9 years (74.5% of all tracts). Standard errors were less than 4.9 years for 89.5% of the tracts, and 95.1% of tracts had standard errors below 5.9 years. Thirteen tracts (2.2%) had standard errors between 7.0 and 9.0 years (see Figure 1). The Virginia Department of Health did not provide standard errors.

HEALTHY PLACES INDEX

The HPI was designed to provide an overall score, ranging from zero (lowest opportunity for health) to 100 (most opportunity for health), for each census tract in the region. The overall HPI score summarizes the independent contribution of 48 census-tract level indicators organized by the six domains.

By design, the HPI score characterizes places (e.g., neighborhoods), not individuals. That is, it describes the influence of neighborhood conditions *irrespective of individual characteristics*. Individual behaviors (e.g., smoking, diet, and exercise) and characteristics (e.g., income, education) are important contributors to health, but place matters to health in ways that transcend individual factors. Tools like the HPI provide metrics about place that can inform policies and actions at the community level, and that can improve health above and beyond the factors that individuals and families can control.

The HPI is useful to anyone interested in learning how local neighborhood conditions influence the health of communities. It was designed to be used by state and local governments, community organizations, health care providers and health systems, public health officials, businesses, and financial institutions. For example, the HPI can be used for:

- Prioritization of investments, resources, and programming in neighborhoods where health needs are the greatest
- Program planning and service delivery
- Community profiles and needs assessments
- Understanding community needs (in conjunction with resident experience)
- Research
- Providing data for grant applications

Domains and Indicators

The 48 indicators were selected based on the following criteria: 1) published research on their association with life expectancy, 2) data quality, 3) and data availability at the census tract level. Based on the preference of the HOC for more recent data and the variation in years involved in the life expectancy calculations, two time periods were chosen for indicators. For example, most Maryland and District of Columbia indicators were drawn from the American Community Survey (ACS) 2014 five-year estimates (2010-2014), while ACS five-year 2013 estimates (2009-2013) were used for Virginia. These indicator years were selected to be recent, as consistent as possible across the region, and within the range of years used to compute life expectancy for that locality.

Based on research literature, expert opinion, and consultation with the HOC, the indicators were then grouped into six “policy action” domains: *Air Quality*, *Education*, *Economic/Other Household Resources*, *Health Care Access*, *Housing*, and *Transportation*. The indicators were standardized to z-scores and scaled in the same direction as life expectancy. A full list of domains, indicators, and data sources is in Table 1. Bivariate correlations between each HPI indicator and life expectancy are shown in Table 3.

Although the racial/ethnic composition and birthplace of residents (and year of entry into the country by immigrants) were not part of the HPI computation, these data (see Table 2) were utilized in a series of additional analyses to estimate the extent to which the segregation of opportunity by race-ethnicity and the influence of racism and discrimination predicted life expectancy (see below). They were not part of the core HPI computations because the focus of the study was on potentially modifiable policy action domains that can improve heath. Teasing apart the complex relationship between race/ethnicity and health through both direct effects (e.g., chronic stress and allostatic load due to exposure to discrimination and trauma) and indirect influences (living in neighborhoods that offer lower opportunities for education, income, housing, and other resources) was beyond the scope of this project. However, a race/ethnicity domain was included in a separate model to assess the impact of race/ethnicity on the other domains in the HPI.

Modeling Methods

The HPI was computed using weighted quantile sum (WQS) regression[4] to estimate life expectancy. The WQS regression method was adopted to accommodate highly correlated data that create collinearity issues and thereby make traditional regression methods problematic. The VCU biostatistician, who had previously used WQS regression models to model colorectal cancer screening adherence[5] and elevated blood lead levels[6] based on neighborhood social determinants of health, adapted that approach for this project.

The objective in this research was to model a health outcome, life expectancy, in relation to a large number of indicators in multiple domains. WQS was used to model life expectancy for an observation (i.e., census tract) as $Y = \beta_0 + \beta_1(\sum_{i=1}^{c} w_i q_i)$, where w_i was the weight parameter for the ith indicator with quantile score q_i, and β_1 was the effect for the combination of indicators. Quantiles were used instead of observed values to reduce the effect of outliers and to account for different scaling of indicators (e.g., median household income and percent with a college degree). Although any reasonable definition of quantiles could be used, including deciles, here quartiles were used to compute the HPI. In the basic WQS model, one weighted index used c number of indicators. The weight w_i represented the relative importance of the indicator and was constrained to range between 0 and 1 and to sum to 1. The weights in the index were estimated through nonlinear optimization using *B* number of bootstrap samples from the data to form the weighted quantile sum index $WQS = \sum_{i=1}^{c} \bar{w}_i q_i$. The final weights in the weighted quantile sum were calculated from the bootstrap sample estimates as $\bar{w}_i = \frac{1}{B}\sum_{b=1}^{B} w_{i(b)} f(\hat{\beta}_{1(b)})$, where $f(\hat{\beta}_{1(b)})$ is a pre-specified signal function. (An example of a signal function is a simple mean function.) The WQS index score was then calculated for each observation in the data set and the overall indicator effect (β_1) was estimated using the model $Y = \beta_0 + \beta_1 WQS$.

The basic WQS model assumed that all indicators in the weighted index shared the same direction of association as the outcome. No two indicators could have different directions of association with the outcome in the WQS model. A negative association for one indicator and a positive association for another indicator is not possible with the WQS model as previously defined. This is due to the shared regression coefficient (β_1) for all the indicators in the index. Thus, prior to modeling, all indicators were transformed to reflect a positive correlation with life expectancy (e.g., percent living in poverty was transformed into the percent *not* living in poverty). Some indicators could have no association (weights estimated to be zero). In addition, some indicators naturally grouped together to form domains, such as education, health insurance, and poverty, and these domains could have different strengths of association with the outcome. To better model multiple sets of diverse indicators, a version of WQS regression called grouped WQS (GWQS) was developed to accommodate multiple domains of explanatory variables with potentially different directions and magnitudes of association with a health outcome. The grouped WQS model included a weighted index and associated effect for each domain of interest, such as $Y = \beta_0 + \beta_1\left(\sum_{i=1}^{c_1} w_i q_i\right) + \beta_2\left(\sum_{j=1}^{c_2} w_j q_j\right) + \beta_3\left(\sum_{k=1}^{c_3} w_k q_k\right) + \beta_4\left(\sum_{l=1}^{c_4} w_l q_l\right) + \beta_5\left(\sum_{m=1}^{c_5} w_m q_m\right) + \beta_6\left(\sum_{m=1}^{c_6} w_m q_m\right)$, where six weighted indexes each had their own effect. As previously noted, the domains represented in the HPI model were Air Quality, Education, Economic/Other Household Resources, Health Care Access, Housing, and Transportation. Each weighted index included a variable number of indicators, and the weights within each index were constrained to sum to 1.

The parameters in the models described above, including the index weights, could be estimated using nonlinear optimization. Simultaneous estimation of the unknown index weights and regression coefficients was achieved through the use of an optimization algorithm that maximized a nonlinear objective function, subject to the linear constraint that the weights in each index summed to 1 and that the weights fell within the bounds [0,1].

The index weights were generated based on data for the entire metropolitan Washington region. Scores for each census tract were determined by application of these weights to the observed indicators in each tract. To create a more interpretable index, the overall GWQS index was rescaled to a range of 0-100. The individual domain indices were also rescaled to sum to the total index in the range 0-100. With this scaling, any domain index could be interpreted as the contribution to the overall index.

Geographic footprint

This project focused on the tract boundaries defined in the 2010 Decennial Census for each jurisdiction that was part of the Metropolitan Washington Council of Governments. This included:

- The District of Columbia
- Maryland
    - Charles County, Frederick County, Montgomery County, Price George County (including Bladensburg, Bowie, College Park, and Greenbelt), Frederick County (including the City of Frederick)
- Virginia
    - Arlington County, Alexandria City, Fairfax County (including Fairfax City and Falls Church), Loudoun County, and Prince William County (including Manassas City and Manassas Park City)

RESULTS

The indicator data for all census tracts in this geographic region were provided to the HOC as a separate Excel file. See Table 1 for full definitions of the indicators.

The HPI results demonstrated that education (33.8% of the HPI weight), economic/other household resources (25.7%), and housing (16.1%) domains exerted the largest influence on the HPI scores (see Table 4). Table 5 displays the weights for individual indicators. As already noted, indicator weights in each domain were rescaled to sum to the total domain weight. Indicators with weights of zero did not influence the HPI score, which means that *in the presence of all other indicators in the model* these indicators had no association with life expectancy. In most cases, this was due to high correlations among one or more of the other indicators rather than the absence of a bivariate association with life expectancy.

See the main report for maps that display the geographic distribution of the HPI scores and a discussion of the policy implications. The HPI was highly correlated ($r=0.77$; $r^2= 0.59$) with life expectancy (see Figure 2). This was a large association, given that life expectancy was modeled solely with tract-level predictors—no individual characteristics (e.g., personal behaviors, household income) were included. Further, not all neighborhood characteristics that influence health (e.g., quality of health care, exposure to racism, etc.) were available for analysis at the census tract level. Thus, the HPI serves as a useful starting place for understanding the health of communities and generating additional questions about other factors. Exploring additional data sources, such as local administrative data and vital statistics, and having conversations with stakeholders and residents can provide insight into other local conditions and specific health outcomes that were not available for analysis at the census tract level.

The domain weights changed when the race/ethnicity domain was added to the analysis, highlighting the degree to which differential exposures experienced by racial and ethnic groups and immigrants impact census tract variations in life expectancy (see Figure 3). The race/ethnicity domain accounted for nearly half (47.0%) of the HPI weight when added to the model, markedly reducing the other domain weights. The greatest influence was on the transportation domain, which exhibited an 85% reduction (from 10.0% to 1.5%) when the race/ethnicity domain was included (see Table 6). It also strongly influenced access to housing (67.1% reduction) and economic/other household resources (47.9% reduction). Education and air quality domains were also reduced by more than a third when the race/ethnicity domain was included in the HPI model (see Figure 3). Table 7 shows how HPI quartiles differed by race-ethnicity and immigrant status. Further work is needed to explore this complex issue and tease apart the various components of the race/ethnicity domain (e.g., immigrant status, race, ethnicity, etc.). However, these initial analyses underscore the added value of applying an equity lens when examining policies that improve population health.

LIMITATIONS

Despite the comprehensiveness of this index and its important contribution to knowledge about neighborhood conditions in metropolitan Washington that contribute to health, there are some limitations of the data that must be noted.

Regional variation

The HPI was developed as a regional tool, optimized to predict variation at the census tract level across the entire metropolitan Washington region. Scores for each census tract were determined by application of these weights to the observed indicators in each tract. It is possible that local dynamics could result in the HPI being more or less predictive in certain areas. For example, the HPI optimized the correlation with life expectancy across the region (r=0.77), but the strength of this correlation varied by jurisdiction (r=0.84 in the District of Columbia; r=0.77 in the Maryland jurisdictions; and r=0.46 in the Northern Virginia localities). Methodological work is ongoing at VCU to develop a spatial version of the GWQS that would take spatial clustering into account and create weights for each location to address this issue in future versions of the HPI.

Life expectancy limitations

There are known limitations to life expectancy estimates, particularly when calculated for small geographic areas.[7] Small area estimates often require aggregating deaths over long time periods, and thus significant changes in the population age distribution of a census tract during these years may bias results. Small death counts at the census-tract level result in life expectancy rates with large standard errors. As previously discussed and shown in Figure 1, 22% of the census tracts in Maryland and the District of Columbia had standard errors greater than or equal to 4.0. High standard errors decrease the reliability of life expectancy estimates. This study did not exclude values with higher standard errors because it was important to provide the HPI for a broader portion of the metropolitan Washington region. Limiting the HPI only to those census tracts with narrow standard errors for life expectancy would defeat the purpose of the project. A comprehensive and inclusive approach was necessary not only to paint a broad picture for the entire region but also to take the first step toward understanding which domains influence variation in life expectancy, mapping community strengths and challenges, and informing policy priorities. Life expectancy data will vary based on the geographic unit (e.g., census tract, zip code, county) and years included in the calculation, so these data may differ from other calculations that are based on a different geography or time period.

Indicator limitations

While the indicator list used in the computation of the HPI is comprehensive and evidence-based, data for many important indicators were unavailable for census tracts, or any sub-county levels. For example, data about social capital, racism, crime, mental health status, drug use, and quality of health care were not available at the census tract level. As a result, the validity of domain weights in quantifying the true importance of a domain in predicting life expectancy must be interpreted with caution. As noted earlier, the proportions, or weights, measure the unique contribution to health *among the factors included in the HPI model*, not *all* of the potential factors that shape health.

Additionally, some of the indicators that were available, such as the air quality measures from the Environmental Protection Agency's National Air Toxics Assessment (NATA) data, are modeled estimates designed as a screening tool, which lack the precision in small areas that exist for other directly measured indicators from the census and other data sources. Low food access among low income populations reflected data based on the 2010 census because that year provided the most recent data available for this indicator. Some indicators may be less relevant in this region. For example, the age of housing—which is typically associated with the risk of exposure to lead-based paint, mold, and other toxins—may not be as relevant in a region where affluent residents occupy historic properties dating as far back as the Colonial era. Some indicators in the HPI may also be markers for other factors not included in the HPI. For example, the correlation observed between limited English proficiency and life expectancy (r=0.24) was likely due to the better health status generally observed among Hispanics (the "Hispanic paradox"[8]) and recent U.S. immigrants[9] and less to the health benefits of language literacy.

Both the indicators and life expectancy rates provide only a snapshot of the population living in the tract at that time. They cannot capture the environment to which the local population was exposed over a lifetime. People who die in one census tract may have lived elsewhere, so the exact conditions to which they were exposed to during a lifetime are difficult to measure. The potential for "mismatch" between indicator data and life expectancy is increased in a highly fluid and rapidly growing region like metropolitan Washington. This growth makes it likely that the characterization of some census tracts in the region based on indicator data that are several years old will not reflect current health opportunities and status.

SUMMARY

The HPI was developed with the understanding that complex systems shaping health are at work at the community level. It therefore may not explain all of the factors that impact life expectancy in all census tracts within the region. Nevertheless, these systems are strongly influenced by many of the opportunities for good health that are captured in the HPI. The HPI is a helpful tool that allows communities to begin to understand the complex interactions between the environment and health outcomes so that action steps can be taken to improve the health of individuals, families, and their neighborhoods. Specific policy recommendations are provided in the main report.

References

[1] County Health Rankings. County Health Rankings & Roadmaps. http://www.countyhealthrankings.org/ Updated March 14, 2018. Accessed July 30 2018.

[2] Metropolitan Washington Council of Governments. *Region Forward: A Comprehensive Guide for Regional Planning and Measuring Progress in the 21st Century Region Forward vision*. https://www.mwcog.org/documents/2010/01/28/region-forward-vision/ Accessed October 15, 2018.

[3] Silcocks PB, Jenner DA, Reza R. Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities. *J Epidemiol Community Health.* 2001;55(1):38–43.

[4] Carrico C, Gennings C, Wheeler D, Factor-Litvak P. Characterization of a weighted quantile sum regression for highly correlated data in a risk analysis setting. *Journal of Agricultural, Biological, and Environmental Statistics*. 2015; 20(1): 100-120.

[5] Wheeler D, Czarnota J, Jones R. Estimating an area-level socioeconomic status index and its association with colonoscopy screening adherence. *PLOS ONE*. 2017;12(6): e0179272.

[6] Wheeler D, Jones R, Schootman M, Nelson E. Explaining variation in elevated blood lead levels among children in Minnesota using neighborhood socioeconomic variables. *Science of the Total Environment*. 2019;650:970-977. https://doi.org/10.1016/j.scitotenv.2018.09.088

[7] Eayres D, Williams ES. Evaluation of methodologies for small area life expectancy estimation. *J Epidemiol Community Health*. 2004 Mar;58(3):243-9.

[8] Ruiz JM, Steffen P, Smith TB. Hispanic mortality paradox: A systematic review and meta-analysis of the longitudinal literature. *American Journal of Public Health*. 2013;103(3):e52-e60.

[9] Antecol H, Bedard K. Unhealthy assimilation: Why do immigrants converge to American health status levels? *Demography.* 2006;43(2):337-60.

Table 1. Metropolitan Washington Healthy Places Index domains, indicators, and data sources.

Domain	Indicator	Description	Data Source*
Air Quality	Cancer risk	Lifetime cancer risk from inhalation of air toxics, per million people.	EPA/NATA[1] (2011)
	Environmental hazards	Exposure to toxins harmful to human health (0-100; higher values = less exposure to toxins)	EPA/NATA[1] (2011)
	Respiratory risk	Ratio of exposure concentration to health-based reference concentration (>1 = increased risk)	EPA/NATA[1] (2011)
Economic/Other Household Resources	Low food access (low income population)	% of low-income population living > 1/2 mile (urban) or >10 miles (rural) from the nearest supermarket, supercenter, or large grocery store.	USDA Food Access Research Atlas (2010)
	Income inequality	Gini Index-measures income distribution among the residents of a specified geography (0 -1; higher values = more inequality)	ACS (table B19083)
	Low food access (overall)	% of total population living > 1/2 mile (urban) or >10 miles (rural) from the nearest supermarket, supercenter, or large grocery store.	USDA Food Access Research Atlas (2013)
	Marital status	% of population 15 yrs+ now married (excluding those who are separated)	ACS (table S1201)
	Median household income	Median annual household income	ACS (table S1903)
	Low income (adult)	% of population ages 18 to 64 yrs with household incomes at or below twice the poverty level (200%)	ACS (table S1701)
	Poverty (adult)	% of population ages 18 to 64 yrs with household incomes below the poverty level (100%)	ACS (table S1701)
	Poverty (child)	% of population under 18 yrs living below the poverty level (100%)	ACS (table S1701)
	Public assistance	% of households receiving public assistance income	ACS (table B19058)
	Single-parent households	% of Children living in households headed by a single-parent	ACS (table B09005)

	Unemployment rate	Percentage of population aged 25-64 who are unemployed	ACS (table S2301)
Education	Preschool enrollment	% of 3 and 4-year-olds not enrolled in school	ACS (table S1401)
	High school diploma/higher	% with a high school diploma/higher	ACS (table S1501)
	Lack of English proficiency	% of households where no one age 14 and over speaks English only or speaks English "very well"	ACS (table S1602)
	Some college/higher	% with some college education or higher	ACS (table S1501)
Health Care Access	Primary care provider access	Ratio of population to PCP's (IM, PA, FM, NP)	HRSA Data Warehouse
	Mental health provider access	Number of mental health providers (county level) divided by the total population per 100,000	County Health Rankings and Roadmaps
	OB/GYN provider access	Ratio of population to OB/GYNs	HRSA Data Warehouse
	Private insurance	% of civilian noninstitutionalized population with private insurance	ACS (table S2703)
	Public Insurance	% of civilian noninstitutionalized population with public insurance	ACS (table S2703)
	Uninsured adults	% of civilian noninstitutionalized population who are uninsured	ACS (table S2701)
	Uninsured children	% of children who are uninsured	ACS (table S2701)
Housing	Older age of housing	Percentage of housing units built 1950 or earlier	ACS (table DP04)
	Overcrowding	% of households with more than one occupant per room	ACS (table DP04)
	Housing vacancies	% of housing units that are vacant	ACS (table DP04)
	Renter occupied	Percentage of occupied housing units not occupied by property owners.	ACS (table DP04)
	Median home value	Median home value of owner-occupied units	ACS (table DP04)
	Median rent	Median rent	ACS (table DP04)

	Housing cost burdened (overall)	% of all households (renters and homeowners) paying more than 30% of income on housing	ACS (tables B25070; B25091)
	Renter housing cost burdened	% of renter households paying more than 30% of income on housing	ACS (table B25070)
	Housing cost burdened (homeowners)	% of owner households paying more than 30% of income on housing	ACS (table B25091)
	Extremely housing cost burdened (overall)	% of all households (renters and homeowners) paying more than 50% of income on housing	ACS (tables B25070; B25091)
	Extremely housing cost burdened (renters)	% of renter households paying more than 50% of income on housing	ACS (table B25070)
	Extremely housing cost burdened (homeowners)	% of owner households paying more than 50% of income on housing	ACS (table B25091)
	Poor housing conditions (renters)	% of rental properties with 1 or more poor housing conditions (e.g., no plumbing, no kitchen, overcrowded, cost burdened)	ACS (table B25123)
	Poor housing conditions (homeowners)	% of homes with 1 or more poor housing conditions (e.g., no plumbing, no kitchen, overcrowded, cost burdened)	ACS (table B25123)
	Housing stability	% of population in the same residence within the past 12 months	ACS (table B07003)
	Housing moves	% of population who moved within the same county within the past 12 months	ACS (table B07003)
	Housing opportunity index	Potential opportunity for Housing Choice Voucher holders seeking housing, Higher values = higher opportunity	Policy Map (2011 data)

Transportation	Commute by motor vehicle	% of population who take a car, taxi, or motorcycle to work	ACS (table S0801)
	Commute by public transit	%of population who take public transport (bus, train, subway) to work	ACS (table S0801)
	Commute by walking/cycling	% of population who walk or bike to work	ACS (table S0801)
	Travel time to work	Average travel time to work (min)	ACS (table S0801)
	Transportation Cost	Low Transportation Cost Index (high value = lower cost)	Policy map (2015 data)
	No access to vehicle	% of households with no access to a vehicle	ACS (table DP04)

[1] EPA/NATA = Environmental Protection Agency National Air Toxics Assessment

Table 2. Immigrant, racial and ethnic Indicators used in additional analyses.

Indicator	Description	Data Source
Asian population	% of the population that is Asian	ACS (table DP05)
Hispanic population	% of the population that is Hispanic	ACS (table DP05)
Multi-race population	% of the population that is multi-race	ACS (table DP05)
NH black population	% of the population that is non-Hispanic black	ACS (table DP05)
NH white population	% of the population that is non-Hispanic white	ACS (table DP05)
Foreign born population	% of population who are immigrants or foreign born	ACS (table B05002)
Immigration pre-1990	% of immigrants who entered the US pre 1990	ACS (table B05005)
Immigration 1990s	% of immigrants who entered the US in the 1990s	ACS (table B05005)
Immigration 2000s	% of immigrants who entered the US in the 2000s	ACS (table B05005)
Immigration 2010+	% of immigrants who entered the US 2010 or later	ACS (table B05005)
African immigrants	% of immigrants who are from Africa	ACS (table B05006)
Asian immigrants	% of immigrants who are from Asia	ACS (table B05006)
European immigrants	% of immigrants who are from Europe	ACS (table B05006)
Latin American immigrants	% of immigrants who are from Latin America	ACS (table B05006)
Diversity Index	Probability that two individuals chosen at random would be of a different race/ethnicity	Policy Map (2011-2015 data)
Racial segregation	Theil Index (0 to 1; higher values = more segregation)	Policy Map (2010 data)

*ACS = American Community Survey 5-year estimates (2014 for D.C. and MD, 2013 for VA)

Table 3. Bivariate correlations with life expectancy at birth by domain.

Domain	Indicator	r
Air Quality	Cancer risk	0.10
	Environmental hazards	0.05
	Respiratory risk	0.01
Economic/Other Household Resources	Public assistance	-0.59
	Single-parent households	-0.54
	Unemployment rate	-0.52
	Median household income	0.52
	Marital status	0.49
	Low income (adult)	-0.47
	Poverty (child)	-0.40
	Poverty (adult)	-0.39
	Low food access (low income population)	-0.23
	Income inequality	-0.01
	Low food access (overall)	0.00
Education	Some college/higher	0.66
	High school diploma/higher	0.48
	Lack of English proficiency	0.24
	Preschool enrollment	-0.18
Health Care Access	Private insurance	0.50
	Public Insurance	-0.47
	OB/GYN provider access	0.33
	Primary care provider access	0.29
	Uninsured adults	-0.22
	Mental health provider access	0.05
	Uninsured children	0.00
Housing	Median home value	0.54
	Median rent	0.51
	Housing cost burdened (overall)	-0.40
	Poor housing conditions (homeowners)	-0.38
	Housing cost burdened (homeowners)	-0.37
	Extremely housing cost burdened (overall)	-0.31
	Housing vacancies	-0.29
	Housing opportunity index	0.27
	Extremely housing cost burdened (homeowners)	-0.25
	Older age of housing	-0.24
	Poor housing conditions (renters)	-0.20
	Renter occupied	-0.19
	Renter housing cost burdened	-0.18
	Housing moves	-0.18

	Overcrowding	-0.18
	Extremely housing cost burdened (renters)	-0.15
	Housing stability	0.02
Transportation	Travel time to work	-0.32
	No access to vehicle	-0.27
	Commute by public transit	-0.20
	Commute by motor vehicle	0.16
	Commute by walking/cycling	0.07
	Transportation Cost	-0.06

Table 4. Metropolitan Washington Healthy Places Index domain weights.

Domain	Weight
Education	0.338
Economic/Other Household Resources	0.257
Housing	0.161
Transportation	0.100
Air Quality	0.088
Health Care Access	0.056
Total	1.000

Table 5. Metropolitan Washington Healthy Places Index indicator weights by domain.

Domain	Indicator	Weight
Air Quality	Cancer risk	0.988
	Environmental hazards	0.012
	Respiratory risk	0.000
Economic/Other Household Resources	Public assistance	0.331
	Marital status	0.242
	Single-parent households	0.184
	Unemployment rate	0.102
	Low food access (overall)	0.092
	Poverty (child)	0.030
	Poverty (adult)	0.012
	Low food access (low income population)	0.003
	Median household income	0.002
	Low income (adult)	0.001
	Income inequality	0.000
Education	Lack of English proficiency	0.484
	Some college/higher	0.275
	High school diploma/higher	0.218
	Preschool enrollment	0.023
Health Care Access	Primary care provider access	0.614
	Private insurance	0.232
	OB/GYN provider access	0.097
	Mental health provider access	0.042
	Uninsured children	0.007
	Uninsured adults	0.004
	Public Insurance	0.004
Housing	Older age of housing	0.238
	Median rent	0.187
	Housing moves	0.137
	Housing vacancies	0.133
	Housing stability	0.108
	Median home value	0.104
	Housing opportunity index	0.034
	Poor housing conditions (homeowners)	0.018
	Overcrowding	0.017
	Renter occupied	0.010
	Housing cost burdened (homeowners)	0.006
	Housing cost burdened (overall)	0.002
	Renter housing cost burdened	0.002
	Extremely housing cost burdened (homeowners)	0.002

	Poor housing conditions (renters)	0.001
	Extremely housing cost burdened (renters)	0.000
	Extremely housing cost burdened (overall)	0.000
Transportation	Travel time to work	0.420
	Commute by public transit	0.359
	Commute by motor vehicle	0.099
	Commute by walking/cycling	0.082
	Transportation Cost	0.040
	No access to vehicle	0.000

Table 6. Change in the Metropolitan Washington Healthy Places Index (HPI) domain weights with and without the race/ethnicity domain in the model.

Domain	Domain weight as a % of HPI (race/ethnicity domain NOT in the model)	Domain weight as a % of HPI (race/ethnicity domain in the model)[1]	Absolute change in domain weight	% change in domain weight
Transportation	10.0%	1.5%	-8.5%	-85.0%
Housing	16.1%	5.3%	-10.8%	-67.1%
Economic/Other Household Resources	25.7%	13.4%	-12.3%	-47.9%
Education	33.8%	21.5%	-12.3%	-36.4%
Air Quality	8.8%	5.7%	-3.1%	-35.2%
Health Care Access	5.6%	5.5%	-0.1%	-1.8%
Race/Ethnicity	-	47.0%		

[1]The weights in this column reflect changes to the final HPI weights in each domain with the addition of the race/ethnicity domain in the model. These were computed to show the variation when race/ethnicity was added to the analysis (highlighting the degree to which differential exposures experienced by racial and ethnic groups and immigrants impact census tract variations in life expectancy).

Table 7. Race, ethnicity, and immigrant status by Metropolitan Washington Healthy Places Index (HPI) quartile.

HPI quartile	**Non-Hispanic Black (%)**	**Non-Hispanic White (%)**	**Hispanic (%)**	**Foreign Born (%)**
1st (high opportunity)	6.2	64.6	9.5	24.6
2nd	13.9	54.3	15.5	25.8
3rd	29.2	36.9	22.6	28.0
4th (low opportunity)	63.0	19.1	12.7	15.0

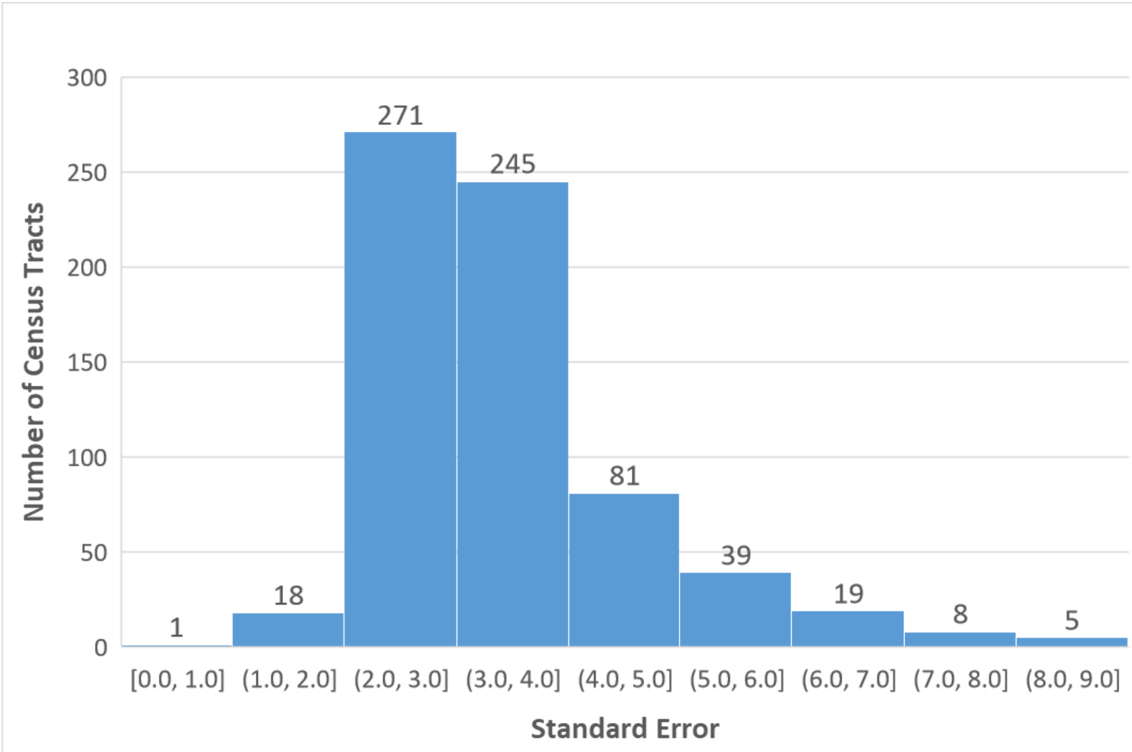

Figure 1. Standard error of life expectancy estimates by census tract, Maryland and the District of Columbia.

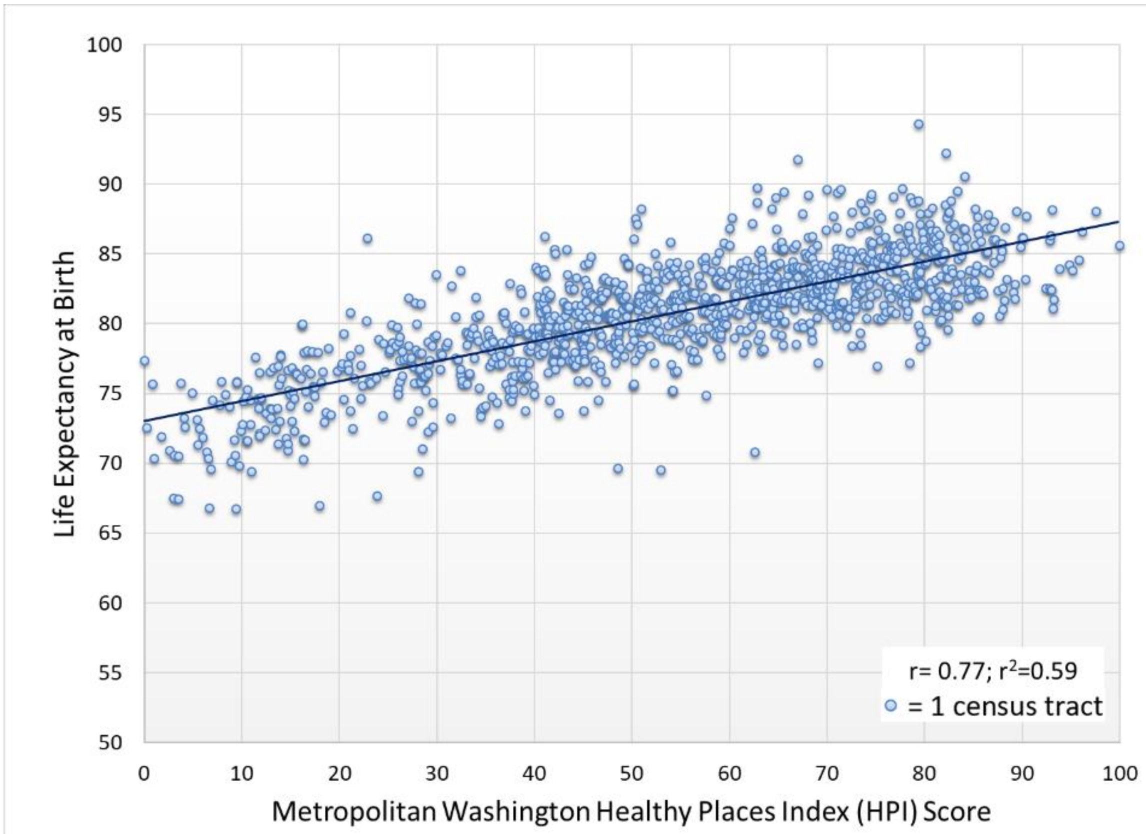

Figure 2. Scatterplot of the Metropolitan Washington Healthy Places Index (HPI) score with life expectancy.

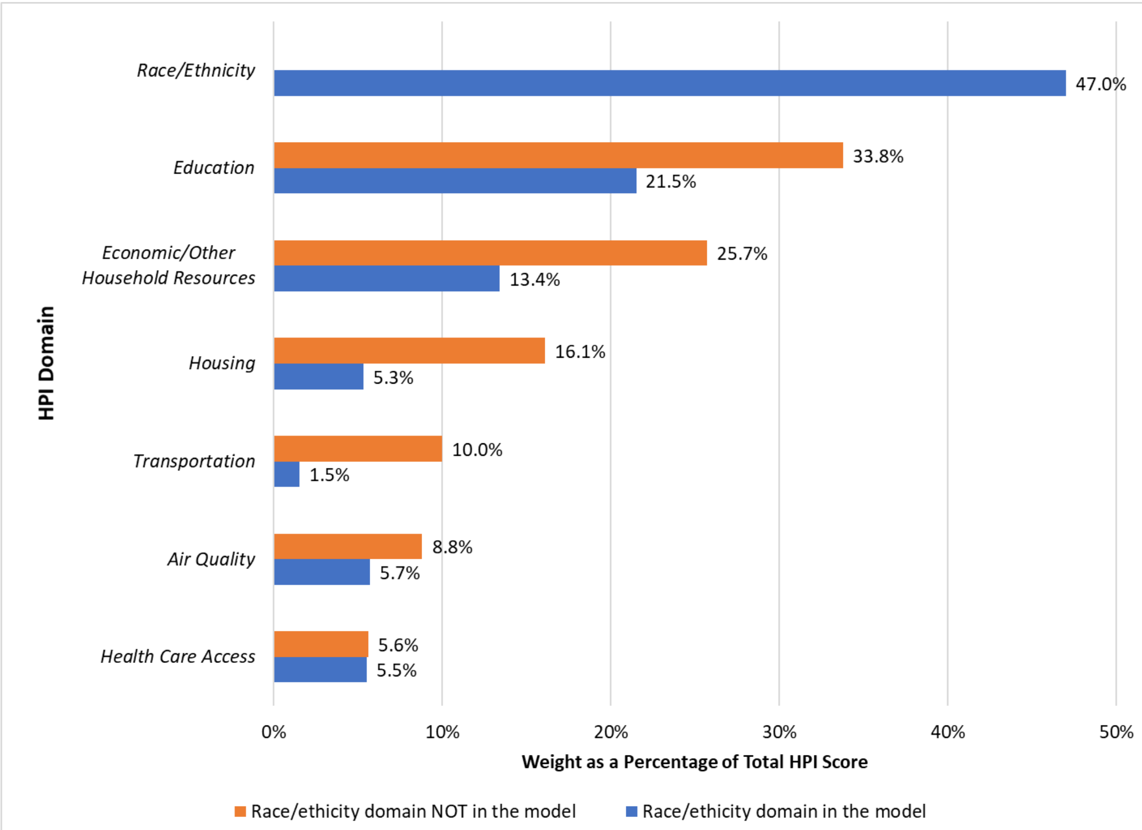

Figure 3. Metropolitan Washington Healthy Places Index (HPI) weights by domain with and without the race/ethnicity domain in the model[1]

[1]The weights in this column reflect changes to the final HPI weights in each domain with the addition of the race/ethnicity domain in the model. These were computed to show the variation when race/ethnicity was added to the analysis (highlighting the degree to which differential exposures experienced by racial and ethnic groups and immigrants impact census tract variations in life expectancy).