Healthy Places Index

http://healthyplacesindex.org

by

Tracy Delaney, PhD
Will Dominie, MPH
Helen Dowling, MPH
Neil Maizlish, PhD, MPH

Public Health Alliance of Southern California
Email: PHASoCal@phi.org
Website: http://phasocal.org/

and

Derek Chapman, PhD
Latoya Hill, MPH
Christine Orndahl, BS
Roy Sabo, PhD
Lauren Snellings, MPH
Sarah Simon Blackburn, MS
Steven Woolf, MD, MPH

Center on Society and Health
Virginia Commonwealth University
Website: http://www.societyhealth.vcu.edu/

July 8, 2018

TABLE OF CONTENTS

TECHNICAL SUMMARY

BACKGROUND: Many governmental entities, academic institutions, and private organizations have developed composite indexes of disadvantage or opportunity. These measures allow policy makers and communities to target interventions and resources to areas with the greatest cumulative extent of deprivation. In 2015, using a framework of the social determinants of health (SDOH), the Public Health Alliance of Southern California ("Alliance") created the Health Disadvantage Index (HDI), which ranked California census tracts on a composite score of disadvantage and displayed the results with maps. The HDI summarized the conditions and the levels of key resources in a community that foster a healthy population and health equity. In 2017, the index was updated in partnership with Virginia Commonwealth University (VCU) Center of Society and Health and a Steering Committee made up of local and state health department representatives and other subject matter experts. The update was also informed by user feedback and included more recent data for indicators, new methods to strengthen the validity and power of the index to predict health outcomes, a version that included a race/ethnicity domain, expanded communications strategies, and enhanced capabilities of the website mapping application. The updated index was rebranded as the California Healthy Places Index (HPI).

INDEX DEVELOPMENT

Data and Methods: The HPI incorporated the same SDOH framing and many of the same domains and indicators as the HDI, whose data generally covered the time period from 2008 to 2012. The HPI is comprised of 25 individual indicators organized in 8 policy action areas (domains) of economy, education, healthcare access, housing, neighborhoods, clean environment, transportation, and social environment. The criteria for including indicators were: 1) publically available data for census tracts with a 2010 population of 1500 or more residents and a group quarters population less than 50% of the total population, 2) evidence from the scientific literature linking the indicator to health, 3) "actionability" through policy, systems, and environmental change, 4) continuity with the HDI, and 5) compatibility with indicator projects sponsored by CDPH. The Steering Committee highlighted a need to validate the HPI against a health outcome and to explore methods to establish domain weights that refined the range of estimates reported in the scientific literature. The Steering Committee also recommended exploring an optional domain for race/ethnicity, residential segregation, and/or racism.

Generally covering the period 2011 to 2015, data from eight primary sources were downloaded from public internet sites, and processed with R and SAS batch programs into individual indicator files. Indicators were screened for a) census tracts with missing data and b) their simple statistical correlation with life expectancy at birth (LEB). A few indicators with a large amount of missing data or correlations with LEB contrary to reports in the literature were excluded from the index. Indicator values were imputed for a small percentage of census tracts with missing data. Indicators in multivariable domains were assessed for co-linearity, which was not present at levels warranting elimination.

Each indicator was scaled in the same direction (higher value $\rightarrow$ more advantage) and standardized using its Z-score. The arithmetic average of individual indicators' Z-scores was calculated for each domain. The HPI total score was the sum of the weighted domain averages. Domain weights were estimated using a regression technique called weighted quantile score, which estimates the relative contributions of correlated variables that tend to cluster in natural groupings (domains), and are associated with an independent outcome. The weights were estimated in such a way to optimize the correlation of the HPI and LEB and its variance-

explained (R square) in simple linear regression. The HPI and domain scores and individual indicator values were also expressed in percentiles so that census tracts could be ranked.

Results: Of California's 8057 census tracts, 7793 met the population threshold and group quarters eligibility criteria. Domain weights were: economy, 0.32; education, 0.19, healthcare access, 0.05; housing, 0.05; neighborhoods, 0.08; clean environment, 0.05; transportation, 0.16; and social environment, 0.10. The census tract distribution of the HPI total score followed a normal (bell shaped) curve. The HPI was highly correlated with LEB (r = 0.56) and explained 31% of the variance (R^2) in simple linear regression. The correlation between HPI score and LEB was diminished slightly in rural areas (r=0.46), but still strong. The overall HPI score and domain scores varied by California region. Counties in the Central Valley, Los Angeles area, and Inland Valley were proportionately more disadvantaged than other California regions. Rural areas tended to have a disproportionate share of the most disadvantaged statewide quartile of census tracts compared to urban areas (28.6% vs. 24.6%). Among several alternative indices of disadvantage (200% of federal poverty level, 80% of median household income, Intercity Hardship Index), CalEnviroScreen 3.0 was most discordant by excluding 651 of the 25% most disadvantaged HPI census tracts (N=1950). Approximately 3.1 million Californians resided in the discordant census tracts. The race/ethnicity version of HPI included a domain whose single indicator was the index of dissimilarity referenced by the county percentage of Black residents. The race/ethnicity domain weight was 0.13 and other 8 domains experienced small reductions in their original weights. The HPI with race had a slightly higher correlation with LEB (r=0.58) and variance explained (R^2 = 0.33) compared to version without race.

COMMUNICATIONS

With feedback from a Communications Committee made up of representatives from the Alliance, VCU, community organizations and CDPH, the HDI underwent significant rebranding as the Healthy Places Index. Indicators and the HPI score were defined using positive language and framing, so that higher numeric values corresponded to greater levels of opportunity or advantage. Two-page briefs for each of 5 distinct audiences were developed to explain the significance of the social determinants of health and how the HPI could be used to prioritize investments to improve the conditions that foster improved community health. The five audiences included health care professionals, state agency policy makers, local government policy makers, business community, and community-based organizations.

The HDI website underwent significant expansion in content and functionality. The HPI score, domains, and individual indicators and their percentile rankings are presented as interactive in interactive maps that provide the values and percentile rankings for 1) all 25 HPI indicators, 8 domains and the overall HPI score and 2) all California for a single indicator, domain, or HPI score. A navigation panel allows users to explore potential policy options for improving health by linking individual domains and indicators to briefs that describe 34 policies and 241 policy actions across the 8 HPI domains. In addition to the 25 HPI indicators, the mapping application also provides 45 selectable data layers covering health outcomes, behavioral risk factors, climate change threats and vulnerabilities, and additional information. The climate change indicators are the same as those used in the CDPH Climate, Health and Equity Program. The mapping application also allows users to pool adjacent census tracts and calculate population-weighted average HPI scores and aggregate census tract data to city, county, and other large geographies. Other features let users create and map their own composite index of individual data layers, upload their own geographies and data, and generate their own community profile report.

BACKGROUND

The purpose of this report is to provide technical information on the Healthy Places Index, HPI, (version 2.0), its 25 constituent indicators, and additional indicators that provide decision-support. The report also provides information on communicating the HPI for different potential audiences and features of the mapping application (http://map.healthyplacesindex.org).

Many governmental entities, academic institutions, and private organizations have developed composite indexes of disadvantage or opportunity.[1-9] These measures allow policy makers and communities to target interventions and resources to areas with the greatest cumulative extent of deprivation. The international practice of disadvantage measurement shares several common concepts and approaches. First, the indexes define deprivation as having multiple dimensions. For example, according to Townsend [10, p.125], people are deprived when they lack the types of diets, clothing, housing, household facilities and fuel and environmental, educational, working and social conditions, activities and facilities which are customary. Second, the experience of disadvantage is a cumulative function of the number and types of deprivation that people experience.[11] Accordingly, deprivation indexes at the small geographic area include the economic resources, social inclusion, health, educational resources, and shared public infrastructure, and physical environmental hazards. Third, the individual domains comprising disadvantage are both components of and consequences of disadvantage. Neighborhood disadvantage predicts poorer human development outcomes, including lower levels of human health, impaired child development, lower educational achievement, and the experience of violence. At the same time, these outcomes may be considered elements of cumulative neighborhood disadvantage.[1-9]

What is the Healthy Places Index?

The Healthy Places Index is the product of the Public Health Alliance of Southern California ("Alliance") who, in 2014, convened a Steering Committee of approximately 20 public health practitioners and researchers from health departments across California (see Appendix A for the list of Steering Committee members), including the California Department of Health and the Bay Area Regional Health Inequities Initiative (BARHII). With Steering Committee guidance, the Alliance staff and consultants conducted literature reviews and embarked on constructing the index. The HPI utilizes the following definition of *health disadvantage*:

> Health disadvantage is the inability of people to fulfill basic human needs required for full social participation and optimal health and well-being. These needs include but not limited to the needs for economic security, food, shelter, safety, transportation, education, social connection and political participation.

The definition incorporates a holistic concept of health and recognition that health is produced by community factors not addressed by our health care system. As articulated by the World Health Organization, health is “a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity” and the fundamental resources for health are “… peace, shelter, education, food, income, a stable ecosystem, sustainable resources, social justice, and equity.” Similarly, the definition of health disadvantage is inclusive of the diverse non-medical economic, cultural, political, and environmental factors that influence physical and cognitive function, behavior, and disease. These factors are often called health determinants or the social determinants of health (SDOH).[12]

Version History and What's New in HPI?

The first version of the HPI was published by the Alliance in February 2015 as the Health Disadvantage Index (1.0).[13] The HDI presented a composite score for each California census tract based on 28 individual indicators, organized into 8 groupings, or domains. The most recent data then available was generally from 2008 to 2012. For each census tract, each indicator was expressed on a standardized scale (Z-scores) of increasing disadvantage, and averaged for each domain. The overall score was calculated as the weighted sum of domain scores. The weights of each domain were informed by a literature review that assessed the relative impact of the domain on overall health status. With enhanced technical documentation, version 1.1 appeared as a Web-based mapping application in January of 2016.[14] In May 2016, the Steering Committee made several recommendations to Alliance staff to guide future versions:

- Update the indicator values using the most recent available data
- Shift a few individual indicators between the education and social domains
- Exclude health outcomes as a component of the index
- Validate the index against a health outcome
- Revise the method for assigning weights
- Explore a version that optionally includes an explicit measure of race/ethnicity, and
- Harmonize the index with indicator projects at the California Department of Public Health.

In 2017, to implement the recommendations, the Alliance partnered with the Center on Society and Health of the Virginia Commonwealth University (https://societyhealth.vcu.edu/) to assist with index construction and help further develop a communications strategy. The result is the Healthy Places Index (HPI) 2.0, which preserves the same conceptual roots as HDI 1.1 from the social determinants of health, but communicates the overall score and indicators in a positive frame – opportunities for community health improvement – rather than as a negative and potentially stigmatizing "disadvantage". The HPI also preserves and updates many of the same domains and indicators as HDI 1.1, but now includes the following features:

- Updated indicator data from the 2011 to 2015 time period
- A revised Education domain consolidating all education indicators
- Health outcomes available to users as decision support layer, but omitted from the index
- Revised domain weights using statistical methods that optimize the correlation between life expectancy at birth and the HPI score for California census tracts
- A separate version of HPI that includes residential, racial segregation as a unique component of the HPI score.

HPI data are the backbone of a robust mapping application (http://map.healthyplacesindex.org) that:

- Visualizes the HPI score and indicator profiles of user-selected census tracts
- Links indicators and domains to actionable policies that address the social determinants of health
- Provides complementary data layers on climate threats and community resilience, health outcomes, health behaviors, and other indices of disadvantage
- Aggregates HPI scores and indicator values to user-selected geographies and jurisdictional boundaries of cities, counties, legislative districts, and other geographies.
- Creates a customized report for user-selected geographies and indicators, and
- Creates a custom index based on the cumulative impact of user-selected indicators.

METHODS: INDEX CONSTRUCTION

Criteria for Domain and Indicator Selection

The grouping of indicators within topic areas or "policy action areas" reflected widely recognized thematic areas of the social determinants of health.[15, 16] The criteria for selection of individual indicators built on those of HDI:

- Accessible public data sources
- Up-to-date data at the geographical level of census tract
- Geographical coverage for all eligible census tracts
- Linkage to policy and other actions ("actionability").

To these, the following were added:
- Association with life expectancy at birth in California census tracts
- Low levels of collinearity with other indicators within a domain
- Continuity with previous versions (HDI 1.1), and
- Alignment with indicator projects at CDPH.

Geographic Inclusion and Exclusion Criteria

In the previous version of HPI, census tracts were included in the index if they had a population of 1500 or greater in the 2010 decennial census AND a group quarters population less than 50% of the total 2010 population. These eligibility criteria aimed to improve the statistical reliability and validity of the index. Census tracts with large share of institutional populations that are mobility restricted (e.g. nursing homes, prisons) and/or are (temporarily) economically dependent on others (e.g. college students) often generate spurious results. The Los Angeles County census tract 6037930401, which was erroneously assigned in the 2010 Decennial Census,[17] was also excluded from the HPI score.

Domains and Indicators

Policy Action Areas (Domains)

We assigned potential indicators to eight thematic groups or policy action areas:

- Economics
- Education
- Healthcare access
- Housing
- Neighborhood conditions
- Clean Environment
- Social environment, and
- Transportation.

Thematic groupings were called "domains" in the previous version of HPI and were modified in several important ways. First, the health domain in HDI 1.1 was discontinued as part of HPI. The health indicators included prevalence of low birth weight, rate of emergency room visits for asthma, prevalence of disabilities, and life expectancy at birth (called years of life lost per capita). The exclusion was necessitated by the proposed domain weighting methodology (see

below), which calculates and optimizes domain weights from the association of index variables and life expectancy at birth (LEB). Inclusion of life expectancy at birth – a health outcome – or other health outcomes as an explicit part of the index would create circular reasoning and blur the distinction between health promoting community conditions (SDOH inputs) from expected health outcomes (output).

The Economics, Education, Social and Neighborhood domains of the HPI are similar to their HDI counterparts. However, the Steering Committee recommended separate domains for housing, transportation, and health care access. The former two reflected the Steering Committee's desire to elevate the importance and distinctiveness of policy actions in transportation and housing. While health insurance was an indicator in HDI's Economic domain, other indicator projects suggested a distinct role for healthcare access and clinical services in the social determinants of health.[18]

Indicator Selection

Applying indicator selection criteria (above), the pool of candidate indicators included existing HDI 1.1 indicators, and new indicators based on recommendations from the Steering Committee and the VCU team. As in HDI 1.0, data at the census tract were not publically available for several desirable indicators including crime, school quality and performance, and health care quality.

To assess geographic coverage, the number of missing census tracts was calculated for each indicator. To assess the association of the indicator with LEB, Pearson correlation coefficients were calculated. For three potential indicators – health insurance, educational attainment of adults, and poverty level – small changes in indicator definition were explored with statistical methods to assess their strength of association with LEB. For health insurance, the question was whether two age-stratified indicators were useful: children (ages 0-17 years) or adults of working ages (18 - 64 years). For educational attainment the question was whether high school completion or a bachelor's degree was more strongly associated with LEB. For poverty, the question was whether 100% or 200% was a stronger predictor of LEB.

Multicollinearity was assessed for each of the domains by calculating a within-domain variance inflation factor (VIF) for each candidate indicator. A VIF of 4 or greater was used as a criteria for identifying excessive multi-colinearity among domain indicators.

Missing Data

As in HDI 1.1, it was anticipated that handful of indicators would have a small percentage of eligible census tracts with missing data. Rather than exclude the entire census tract from the HPI, imputation of missing data was done using an algorithm that took into account the spatial distribution of data in neighboring census tracts (knnImputation option in the DMwR R package).

Indicator Standardization and Scaling

Each indicator was standardized by computing its Z-score, which is aligned so that higher values indicated greater advantage. This required "flipping" (multiplying by -1 or subtracting from 100%) for Clean Environment variables and severe housing cost burden, which were framed in the negative direction: higher values indicate less advantage.

For a given indicator, the Z score, Z, for the *i*th census tract is the difference between the census tract value, *X*, and the overall variable mean, μ, divided by the variable's standard deviation, σ:

$$Z_i = \frac{X_i - \mu}{\sigma}.$$

Weights

Domain Weighting

Domain weights were empirically estimated using weighted and constrained least squares regression model of the eight domain scores against LEB.[19] This model is fit using the eight domain scores and LEB for each census tract. This model simultaneously estimates the domain weights and the association between HPI 2.0 and LEB in such a way that:

- The association between HPI 2.0 and LEB is maximized.
- Domains are allotted more weight if they contribute more the prediction of LEB.
- All Domains are guaranteed a minimum 5% weight. Because all domains and indicators were carefully chosen based on expert opinion and evidence for an association with health in the literature, we wanted to ensure that each domain retained a minimum weight. A minimum weight of 5% leaves 60% of the weighting to the modeling process (8 domains $\times$ 5% = 40% will be in the model based on this criteria).
- If a modeled domain weight were much larger than expected based on expectations from the literature and prior experience with an index like this, we would consider instituting an upper bound with the advice and consent of the Steering Committee. (This contingency did not occur.)

Missing data were excluded in the initial calculation of domain weights. For comparison in a subsequent sensitivity analysis, domain weights were also calculated using imputed values for missing data.

Final Index and Percentile Rankings

The Z-scores of each domain were first averaged to produce a domain score, $\bar{Z}$. The HPI was then calculated for each census tract by multiplying each domain score by the corresponding estimated domain weight, and summing across the eight domains.

$$\text{HPI} = (w_1 \times \bar{Z}_{Economic}) + (w_2 \times \bar{Z}_{Education}) + (w_3 \times \bar{Z}_{HealthcareAccess}) + (w_4 \times \bar{Z}_{Housing}) + (w_5 \times \bar{Z}_{Neighborrhods}) + (w_6 \times \bar{Z}_{Clean\ Environment}) + (w_7 \times \bar{Z}_{Social}) + (w_8 \times \bar{Z}_{Transportation})$$

The census tract percentile of individual indicators, domain $\bar{Z}$ scores, and the overall HPI score was based on their rank order among 7793 census tracts. Ties were assigned the arithmetic average of their ranks. Zero percentile was most disadvantaged and 100 percentile was the most advantaged. In the case of high school enrollment, which had a large percentage of census tracts with 100% enrollment, the percentile of these ties (63.8 percentile) were recoded as 100 percentile. This was done to place the census tracts with the best outcome in the highest quartile for the purpose of mapping the indicator. Similarly, for wildfire risks (a decision support indicator), a large percentage of the census tracts were not in high severity zones. The percentile ranking (61.4 percentile) was recoded to 100 percentile so map legend color would show the best outcomes for these census tracts.

Validation of Domain Weights

To gauge the sensitivity of the domain weights to our estimation methods, we performed four sensitivity analyses. These included estimating domain weights

(i) after imputing missing data
(ii) using a machine learning approach
(iii) using a purely algorithmic approach, and
(iv) accounting for distinctions between rural and urban census tracts.

For *Machine Learning*, we used algorithmic computer programing (rather than probabilistic estimation) to search through all possible combinations of domain scores, selecting that combination of domain weights that led to the largest possible association between HPI and LEB. For the *Purely Algorithmic Approach*, we estimated domain weights for an unconstrained model, excluding any indicators that were not associated with LEB. The domain weights for this model were estimated without constraint (i.e. no 5% minimum so that domain weights can approach 0%).

Rural vs. Urban

The consistency of HPI in rural and urbanized census tracts was analyzed through a comparison of correlations between the index and LEB. The definition of rural and urban followed the three categories used in U.S. 2010 Census and American Community Survey, which factors population thresholds, populations density, land use, and distance to and continuity with adjacent population centers.[20] Generally, urbanized areas are those with 50,000 or more people. Urban clusters are areas with at least 2,500 but fewer than 50,000 people, and rural is any other area.

Race/ethnicity Version of HPI

The HPI does not have an explicit domain or indicator that reflects race/ethnicity. This reflects the desire to make the HPI be accessible to state policy makers, who are limited by the 1996 ballot proposition 209 from making funding decisions for state programs based on race. Race/ethnicity is widely accepted component of the social determinants of health and of health inequity, and the HPI Steering Committee requested that a complementary HPI be created to include some aspect of race/ethnicity.

There are scores of metrics that have been used to describe race, racial attitudes, racism, and residential racial segregation.[21, 22] Some metrics rely on population surveys such as the General Social Survey,[23] which do not have census tract geographical resolution required by the HPI. Using U.S. Census data, the simplest metrics use the population percentage of a specific race/ethnic group. However, there is no agreement on the threshold that constitutes racial segregation (e.g., range 25% to 60%). Of the metrics that describe geospatial patterns of residential housing, the Index of Dissimilarity (IOD) is the most widely used.[24-27] It measures the evenness of the racial distribution in a small geography compared to the metropolitan area, region, or state in which it is a part. It is sometimes described as a measure of the in-migration or out-migration of a specific race/ethnic group from a small geographic area (e.g. census tract) that would achieve racial parity with the larger region. The mathematical formula is given by:

$$\frac{\sum_{i=1}^{n}\left[t_i\left|(p_i - P)\right|\right]}{[2TP(1-P)]}$$

where

i is the geographic unit (e.g. census tracts, census blocks) of aggregation up to *n* units
t_i is the total residents in the *i*th geographic unit
p_i is the proportion of a specific minority group in the *i*th geographic unit
P is the proportion of a specific minority group in the larger geographic reference area
T is the total population in the larger geographic reference area.

The formula was applied to California census tracts using 2010 census blocks[28] as the aggregation unit (i) and the respective county proportion of a race/ethnic group as the reference measure of P. Only census blocks that fell within census tracts meeting HPI inclusion criteria were included. Applying the formula to a specific minority group generates an IOD that is referable only to that group. Several IODs were computed based on different mutually exclusive race/ethnicity groupings: non-Whites, Blacks, Latinos, and Asians. Due to small numbers, IODs were not calculated individually for Native Hawaiians and Other Pacific Islanders, Native Americans/Alaskan Natives, Multiple race, or Other race, although these groups were included in the IOD for non-Whites. Simple Pearson correlation coefficients, r, were computed between each race/ethnicity-specific IOD and life expectancy at birth (LEB). Only the association of life expectancy and the IOD for Blacks demonstrated a negative association (Black: r = -0.33, Asian: r = 0.12, Latino: r = 0.10; non-Whites: r = 0.02). The IOD_{Black} was selected as an individual indicator and the single constituent of a ninth policy action area (domain).

Using the same methods as those used for the HPI (8 domains), a separate HPI (HPI_{Race}) was created in which all nine domain weights were estimated using weighted quantile sums regression.

Decision-Support Indicators and Domains

Although indicators in the HDI 1.1 health domain were no longer part of the HPI, the Steering Committee recognized that geographically refined health outcomes and other data are valuable information that can be used in conjunction with the census tract HPI scores and rankings. This recognition also led to the identification of additional indicators to support decision-making with the HPI. These include:

- Health outcomes from the 500 Cities Project [29, 30] that has substantial geographic coverage of census tracts in California urban areas
- Climate threats
- Demographic and built environment indicators of climate change vulnerability or resilience developed by CDPH, and
- Candidate indicators for the HPI that did not meet all the inclusion criteria.

These indicators were also updated with the most recent data available and comprise an expanded list of "decision-support" indicators that were included in the HPI mapping application.

Comparative and Concordance Analyses

The Healthy Places Index was compared to individual indicators and indices that are used by California governmental agencies and local health departments to define disadvantaged communities (Table 1). These include:

- CalEnviroScreen[31], developed in 2014 by the California Environmental Protection Agency,
- the federal poverty level, a long-standing component of many indices of disadvantage
- 80% of the median household income, and
- (Intercity) Hardship Index, developed by academic researchers 1980s and updated in 2007[32, 33], and used by the Los Angeles Department of Public Health.[34]

Table 1. **Description of Indices to Describe Community Disadvantage**

Index/ Indicator	CES 3.0 [31]	Hardship Index [32, 33]	Poverty/ 80% Median Income
Purpose	Identify pollution- burdened census tracts for enforcement, training, and public outreach	Identify economically disadvantaged communities	Identify economically disadvantaged communities
Concep-tual basis	Disadvantage is determined by exposure to air, soil, and water pollution and its interaction with biological, social, and health characteristics of population subgroups	Economic hardship is a lack of economic resources of working populations and dependency of populations unable to generate economic resources	The amount of family income falls below a threshold to sustain adequate standard of living
Number of Indicators	19	8	Poverty, 200% of federal poverty level; 80% of median household income
Domains	Pollution Burden and Population Burden	Dependency, low education, income-poverty-overcrowding	NA
Standardi zation of indicators	Quintile distribution score (1-5)	Percent change from baseline of the range of values in geographic area	NA
Weighting	Equal	Equal	N/A
Final Score	Multiply domain scores, rescale	Add scores, rescale 0-100	

CalEnviroScreen 3.0

CalEnviroScreen (June 2018 update) was developed by the California Environmental Protection Agency, which states "CalEnviroScreen is primarily designed to assist the Agency in carrying out its environmental justice mission to conduct its activities in a manner that ensures the fair treatment of all Californians, including minority and low-income populations." Its original purpose was to guide its internal enforcement, education, and training efforts. With the passage of SB535, its use has broadened to include designating disadvantaged communities for enhanced funding in California climate change and transportation programs.

CES organizes indicators into two broad domains representing pollution burden and population characteristics, which include sensitive populations and socioeconomic factors. The final score represents the multiplicative interaction of the pollution burden and population characteristics. This follows observations in the epidemiologic and toxicological literature in which population characteristics amplify health effects of specific environmental pollutants. Seven of the 19 indicators in CES are also used in the HPI (unemployment, poverty, low educational attainment, PM2.5, ozone, diesel particulate matter, drinking water contaminants).

Intercity Hardship Index

The Intercity Hardship Index, also known as the Urban Hardship Index and Economic Hardship Index is a composite of 6 indicators:

- Unemployment rate,
- Dependency (percent of the population aged >65 years or < 18 years),
- Low education attainment (the percentage of those over the age of 25 with less than a high school education),
- Per capita income,
- Housing overcrowding (>1 occupant/room),
- Poverty level (<100% of federal poverty).

Except for dependency, the 5 other indicators in the Hardship Index are part of the HPI.

Poverty

Multiples of the federal poverty level are commonly used to describe economic disadvantage and establish eligibility for some federal and state health and human service programs. The poverty level is an income threshold adjusted for family composition and size and includes money income before taxes, but excludes capital gains and noncash benefits such as public housing, Medicaid, and food stamps.[35] The Women, Infant, and Children Program[36] administered by the California Department of Public Health is an example of a state governmental program that uses the federal poverty level (185% of FPL) to establish program eligibility. For HPI and CES, poverty was defined at 200% of the federal poverty level. The Hardship Index uses 100% of the poverty level to define poverty.

Median Household Income

Percent of median household income for a given geographic area is used by several governmental agencies to define low income households that are eligible for benefits programs

such as housing assistance,[37, 38] or a disadvantaged community (80%).[39] Median household income is one of the indicators in the HPI economic resources domain.

Agreement Statistics

To describe the degree of census tract agreement between the different indices, we computed agreement statistics for dichotomous variables in 2 by 2 contingency tables:

		Alternative Index	
HPI		Most Dis-Advantaged 25%	Least Dis-Advantaged 75%
Most Disadvantaged 25%		a	b
Least Disadvantaged 75%		c	d

where Sensitivity = $\frac{a}{a+c}$

Specificity = $\frac{d}{b+d}$

Positive predictive value = $\frac{a}{a+b}$

Proportion of agreement = $\frac{a+d}{a+b+c+d}$;

For HPI, CES, the Hardship Index, and poverty we dichotomized the percentile distribution of the total score at 25% (i.e. 25% most disadvantaged census tracts). The 25% percentile cut point was chosen because California EPA designated the 25% highest scoring census tracts in CES as disadvantaged communities. We chose cut points above and below 80% of the 2011-2015 California annual median household income ($61,818 \times 0.8 = $49,454).

We computed sensitivity, specificity, positive predictive value, and proportion of agreement for the different index comparisons using HPI as the screening variable and the alternative index as the reference. In addition to the number of census tracts, we used 2010 Census populations of the census tracts to estimate the size of residential population in agreement or disagreement areas. All comparisons included only HPI eligible census tracts based on a 2010 population of ≥1500 residents and a group quarters population <50%.

Data Processing and Quality Assurance Procedures

Data were acquired as comma separated values files from public websites of the organizations that developed or processed data from primary sources. R and SAS programs (available by request) were written to abstract numerator, denominator, and outcomes (e.g. percent or rate), and the margin of error when available. The specific construction of indicators from source files is provided in Appendix B (Data Dictionary and Source Data Variable Transformations for HPI 2.0 Files). Data quality was first checked by the VCU team by examining distributions, missing data, and potential outliers of individual indicators and their percentile rakings (for correct directionality). The resulting data files were rechecked by Alliance staff using an R program that generated distributions, missing data, Z-scores, and domain averages, and recomputed the HPI score using reported domain weights. A discrepant indicator was checked and corrected, if necessary, until the indicators values matched exactly or with slight rounding error.

INDEX RESULTS

Census Tract Eligibility

We retained the same eligibility criteria as HDI1.1. Of the 8057 California census tracts in the U.S. 2010 Decennial Census, 7,794 met our eligibility criteria based on population size (≥ 1,500) and living in group quarters (<50%) . Of the 230 excluded census tracts, 142 met both criteria, 25 were excluded because of insufficient population alone, and 63 were excluded for group quarters alone. One Los Angeles County census tract (6037930401) was erroneously assigned in the 2010 Decennial Census[17] and was excluded. Thus, 7,793 census tracts were eligible.

HPI Domains and Indicators

Indicators Selected and Omitted

We obtained data on 38 potential indicators from eight online public sources: U. S. Census Bureau's American Community Survey (ACS), California Environmental Protection Agency (CalEPA), US Department of Housing and Urban Development (HUD), Green Info, (parks), the National Land Cover Database, (tree canopy), US Department of Food and Agriculture (supermarket access), US Environmental Protection Agency (retail density), and University of California, Berkeley (voter participation). Virginia Commonwealth University provided access to their analysis of life expectancy at the California census tract level. Of the 38 candidate indicators, 25 met all the criteria for inclusion in the HPI (Table 2, Figure 1). Missing data and correlations with life expectancy for all 38 candidate indicators are presented in Table 3. The physician ratio (number of physicians per 100,000 population) was missing for a majority of census tracts, and was excluded. (Table 3). A small percentage of census tracts (< 1%) had missing data in the Education domain.

The percentage of civilian noninstitutionalized population 18 to 64 years with health insurance coverage (R^2=0.10) was chosen over percentage of civilian noninstitutionalized population under 18 years with health insurance coverage (R^2=0.035) for the health insurance indicator. The percentage of the population ages 25 years and older that completed a Bachelor's degree or higher (R^2=0.29) was chosen over the percentage of the population ages 25 years and older that completed high school or equivalency (R^2=0.09) for the educational attainment indicator. The percentage of the population living at or above 200% of the federal poverty level (R^2=0.23) was selected over the percentage of the population living at or above 100% of the federal poverty level (R^2=0.20) for the poverty indicator.

Several potential indicators were omitted because their associations with LEB were contrary to expectations based on the scientific literature or well-known, but paradoxical findings in the scientific literature. This includes the "healthy immigrant effect" and "Latino Paradox" [40, 41] that the communications committee believed was challenging to communicate. Among contrary findings, traffic density, low supermarket access using a 1 mile threshold for urban areas and a 10 mile threshold for rural areas, and linguistic isolation (English language proficiency) were positively associated with increasing LEB. Consistent with the Latino or immigrant paradox, the percentage of foreign-born (immigrants) was positively correlated with LEB. Moreover, traffic density, supermarket access, and English language proficiency were all correlated with the percentage of foreign-born. Correlations with these indicated that were stratified by census tracts with high and low percentage of foreign-born, showed correlations that differed on LEB without stratification. This suggests that confounding by foreign-born may play a role in contrary findings (i.e. inconsistent stratum-specific and overall measures of association).

Table 2. Policy Action Areas (Domains), Indicators and their Data Sources for the Healthy Places Index 2.0

Policy Action/ Indicator	Definition	Data Source†, Year
Economic		
abovepoverty*	Percent of the population with an income exceeding 200% of federal poverty level	ACS[42], 2011-2015
employed*	Percentage of population aged 25-64 who are employed	ACS[42], 2011-2015
income*	Median Household Income	ACS[42], 2011-2015
Education		
bachelorsed	Percentage of population over age 25 with a bachelor's education or higher	ACS[42], 2011-2015
inhighschool*	Percentage of 15-17 year olds enrolled in school	ACS[42], 2011-2015
inpreschool*	Percentage of 3 and 4 year olds enrolled in pre-school	ACS[42], 2011-2015
Healthcare Access		
insured*	Percentage of adults aged 18 to 64 years currently insured	ACS[42], 2011-2015
Housing		
homeownership*	Percentage of occupied housing units occupied by property owners	ACS[42], 2011-2015
houserepair	Percent of households with complete kitchen facilities and plumbing	ACS[42], 2011-2015
ownsevere	Percentage of low income homeowners paying more than 50% of income on housing costs	CHAS, 2010-2014
rentsevere	Percentage of low income renter households paying more than 50% of income on housing costs	CHAS, 2010-2014
uncrowded*	Percentage of households with less or equal to 1 occupant per room	ACS[42], 2011-2015
Neighborhood		
parkaccess*	Percentage of the population living within ½ -mile of a park, beach, or open space greater than 1 acre	GreenInfo[43], 2012
treecanopy*	Population-weighted percentage of the census tract area with tree canopy	NLCD[44], 2011
supermkts*	Percentage of the urban and small town population residing less than 1/2 mile from a supermarket/large grocery store, and the percent of the rural population living less than 1 miles from a supermarket/large grocery store	USDA[45], 2015
alcoffsale	Percentage of the population residing within ¼ mile of an off-site sales alcohol outlet	ABC[46], 2014
retail*	Combined employment density for retail, entertainment, and educational uses (jobs/acre)	USEPA[47], 2010

Table 2. Policy Action Areas (Domains), Indicators and their Data Sources for the Healthy Places Index 2.0

Domain/ Indicator	Definition	Data Source, Year
Clean Environment		
dieselpm	Spatial distribution of gridded diesel PM emissions from on-road and non-road sources for a 2012 summer day in July (kg/day) Census tracts were ordered by diesel PM concentration values and assigned a percentile based on the statewide distribution of values.	CalEPA[48], 2012
h20contam	Cal EnviroScreen 3.0 drinking water contaminant index for selected contaminants	CalEPA[48], 2015-2013
ozone	Mean of summer months (May-October) of the daily maximum 8-hour ozone concentration (ppm), averaged over three years (2012 to 2014). Census tracts were ordered by ozone concentration values and assigned a percentile based on the statewide distribution of values.	CalEPA[48], 2012-2014
pm25*	Annual mean concentration of PM2.5 (average of quarterly means, $\mu g/m^3$), over three years (2012 to 2014) Census tracts were ordered by PM2.5 concentration values and assigned a percentile based on the statewide distribution of values.	CalEPA[48], 2012-2014
Social		
voting*	Percentage of registered voters voting in the 2012 general election	UC Berkeley[49], 2012
twoparents*	Percentage of family households with children under 18 with two parents	ACS[42], 2011-2015
Transportation		
automobile*	Percentage of households with access to an automobile	ACS[42], 2011-2015
commute	Percentage of workers (16 years and older) commuting by walking, cycling, or transit (excluding working from home)	ACS[42], 2011-2015

Also included in HDI 1.1

† ABC, Alcoholic Beverage Commission; ACS, American Community Survey; CHAS, Comprehensive Housing Assessment System; CalEPA; California Environmental Protection Agency; NLCD, National Land Cover Database; USDA FARA, U.S. Department of Agriculture Food Access Research Atlas; USEPA, U.S. Environmental Protection Agency; UC Berkeley, University of California, Berkeley

Table 3. Candidate HPI Indicators and Inclusion/Exclusion Criteria

HPI variable name	Domain/Indicator Short Name	Miss-ing N	Correlation with LEB, r	Reason for Being Omitted from HPI
	Economic	2		
employed	Employed	0	0.46	
income	Median Household Income	2	0.52	
abovepoverty	Above Poverty at 200% Level	0	0.49	
	Above Poverty at 100% Level	0	0.45	Weaker association with LEB than 200%
	Gini Index	0	-0.02	Contrary association with LEB
	Education	153		
inhighschool	Enrolled in High School	65	0.07	
inpreschool	Enrolled in Primary School	111	0.28	
bachelorsed	Education of Bachelor's or Higher	0	0.55	
	Education Equal or Greater than High School	0	0.31	Weaker association with LEB than Bach.
	Healthcare Access	4062		
insured	Insured Adult	0	0.33	
	Insured Child	11	0.19	Weaker association with LEB than adult
	Physician Ratio	4055	–	Excessive missing data
	Housing	23		
uncrowded	Uncrowded	0	0.17	
homeownership	Homeownership	0	0.20	
houserepair	Housing with complete kitchen and plumbing	0	0.07	
ownsevere	Low Income Home Owners Severe Cost Burden	23	-0.13	
rentsevere	Low Income Renters Severe Cost Burden	0	-0.24	
	Renting Cost Burden (all income levels)	0	-0.25	Not consistent with HDC criteria
	Owning Cost Burden (all income levels)	0	-0.08	Not consistent with HDC criteria
	Neighborhoods	0		
parkaccess	Park Access	0	0.12	
treecanopy	Tree Canopy	0	0.09	
supermkts	Food/supermarket Access	0	0.02	
retail	Retail Density	0	0.05	
alcoffsale	Off-site Alcohol Outlets	0	0.11	
	Traffic Density	0	0.07	Contrary association with LEB

	Clean Environment	19		
ozone	Ozone	0	-0.24	
pm25	PM 2.5	17	-0.08	
dieselpm	Diesel PM	0	-0.01	
h20contam	Water Contaminants	3	-0.09	
	Pesticides	0	0.02	Contrary association with LEB
	Toxic releases from facilities	0	0.02	Contrary association with LEB
	Social	0		
voting	Voting	0	0.31	
twoparents	Two Parent Household	0	0.38	
	Foreign Born	0	0.18	Difficult to message
	English proficient	0	-0.01	Contrary association with LEB
	Transportation	0		
automobile	Automobile access	0	0.20	
commute	Active Commute	0	0.05	

ECONOMIC 0.32	EDUCATION 0.19	HEALTHCARE 0.05	HOUSING 0.05	NEIGHBOR-HOOD 0.08	CLEAN ENVIRONMENT 0.05	SOCIAL 0.10	TRANSPOR-TATION 0.16
• Poverty • Employment • Income	• Pre-school enrollment • High school enrollment • Bachelors attainment	• Insured adults	Severe cost burden low-income: • renters • owners • Homeownership • Kitchen and plumbing • Crowding	• Retail jobs • Supermarket access • Parks • Tree canopy • Alcohol establishments	• Diesel PM • Ozone • PM2.5 • Drinking Water	• Two Parent Household • Voting	• Healthy Commuting • Automobile access

Figure 1. Health Places Index Policy Action Areas (Domains), Weights, and Individual Indicators

With the final set of 25 indicators, the assessment of multicollinearity within domains did not show any VIF values above 4 and no indicators were excluded within domains due to concerns with multi-colinearity.

Weights, Correlation Between HPI Score and LEB, and R^2

Applying the WQS package in R to HPI indicators (without imputation of missing data), weights were obtained for the eight domains, as shown in Table 4, Column A and Figure 1.

Table 4. Weighted Quantile Sums Domain Weights

		Sensitivity Analysis		
	A	B	C	D
Domain	WQS No Imputation	WQS Imputation	Machine Learning	WQS-Unadjusted
Economic	0.32	0.34	0.40	0.46
Education	0.19	0.18	0.15	0.23
Housing	0.05	0.05	0.05	<0.01
Insurance	0.05	0.06	0.05	<0.01
Neighborhood	0.08	0.09	0.05	0.04
Clean Environment	0.05	0.05	0.05	<0.01
Social	0.10	0.09	0.05	0.09
Transportation	0.16	0.15	0.20	0.18

Note: Machine Learning included a 0.05 minimum weight; $WQS_{unadjusted}$.did not include a minimum weight

Using the weights in Column A above, the correlation between LEB and the HPI score was strong (r = 0.56) and a large proportion of the variation was explained (R^2 = 0.31) in simple linear regression. In none of the sensitivity analyses (B-D), did the estimated weights change meaningfully from our original hybrid approach (Column A).

Rural/Urban

Associations (Pearson r) between life expectancy at birth and the HPI score were positively correlated in each of three strata of urbanization and showed a small decline with increasing levels of rurality: 0.56 (N=7051 urban census tracts), 0.46 (N=384 urban clusters in rural areas), and 0.42 (N= 358 rural census tracts).

Descriptive Analyses

This frequency distribution of census tracts by HPI score is presented in Figure 2. The distribution ranges from -1.96 to 1.5 with a mean centered at 0, and approximates a normal curve.

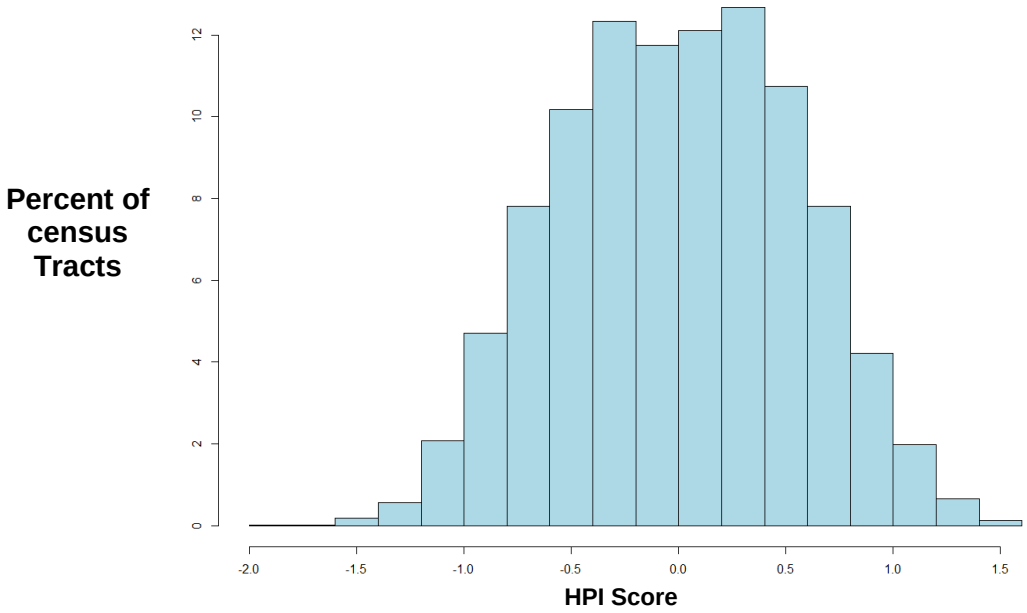


Figure 2. Distribution of HDI scores for 7,793 California census tracts

Most indicator domains are positively correlated with each other. Figure 3 illustrates Pearson correlations among the HPI score and component unweighted indicator domain scores. Domain scores for economic, education, social and healthcare domains tended to have high correlations with each other.

Domain	Economic	Education	Health-care	Housing	Neighbor-hood	Clean Environment	Social
Education	0.62						
Healthcare	0.72	0.56					
Housing	0.65	0.37	0.69				
Neighborhood	0.27	0.26	0.27	0.11			
Clean Environment	0.30	0.25	0.34	0.25	0.28		
Social	0.70	0.46	0.63	0.63	0.23	0.39	
Transportation	0.43	0.34	0.31	0.28	0.14	0.29	0.37

Figure 3. Pearson Correlations Among HPI Domains, California, 2010

Impact of Weighting Domains (compared to equal weighting)

Of the 1948 census tracts comprising the most disadvantaged 25% of HPI scores, 245 (12.6%) were discordant between the HPI and a re-calculated HPI in which there was no weighting (Table 5). Weighting has a modest impact on membership of census tracts in the most disadvantaged quartile.

Table 5. Concordance of HPI Weighted and Equal Weighted Domains

Census Tracts

		HPI 2.0 Most Disadvantaged 25%		
HPI Equal Weight↓		Y	N	Sum
Most Disadvantaged 25%	Y	1703	245	1948
	N	245	5600	5845
	Sum	1948	5845	7793

Geographic Distribution of HPI and Domain Scores

Table 6 gives the distribution of census tracts by quartile of HPI score by California region. The San Joaquin Valley and Inland Empire have a disproportionate share of census tracts in the most disadvantaged quartile and lower mean HPI scores (Table 7). The Bay Area has the smallest share of disadvantaged census tracts. Population counts show a similar pattern (Table 8).

Table 6. Distribution of Disadvantaged Census Tracts by Region, California, 2010

	Quartile of HPI Score					Percent
Region	Least 1	2	3	Most 4	Sum	Disadvantaged (Most/Sum)
Bay Area	821	405	228	94	1548	6
Inland Valley	49	171	259	322	801	40
Los Angeles County	608	643	754	828	2833	29
Sacramento Area	117	162	138	84	501	17
San Diego	174	186	159	123	642	19
San Joaquin Valley	24	128	197	387	736	53
Other	155	253	214	110	732	15
Sum	1948	1948	1949	1948	7793	25

† Regions by County:
Bay Area: Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano, Sonoma
San Joaquin Valley: Fresno, Kern, Kings, Madera, Merced, San Joaquin, Stanislaus, Tulare
Inland Valley: Riverside, San Bernardino
Sacramento Area: El Dorado, Placer, Sacramento, Sutter, Yolo, Yuba
San Diego: Imperial, San Diego
Other: Butte, Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne, Monterey, San Benito, Santa Cruz, Del Norte, Humboldt, Lake, Mendocino, Trinity, Lassen, Modoc, Nevada, Plumas, Sierra, Siskiyou, Colusa, Glenn, Tehama, San Luis Obispo, Santa Barbara, Shasta, Ventura

Table 7. Mean HPI and Domain Scores by California Region

Region	HPI Total	Economic	Education	Health Care	Housing	Neighbor-hood	Clean Environm-ent	Social	Trans-portation
Bay Area	0.42	0.57	0.34	0.63	0.17	0.19	0.59	0.46	0.26
Inland Valley	-0.29	-0.38	-0.38	-0.27	0.09	-0.23	-0.58	-0.21	-0.11
Los Angeles	-0.07	0.00	0.05	-0.31	-0.25	-0.06	-0.23	-0.29	-0.03
Sacramento Area	0.05	-0.08	-0.03	0.36	0.23	0.22	0.30	0.22	-0.04
San Diego	0.06	0.04	0.04	0.09	0.13	-0.04	0.29	0.24	-0.04
San Joaquin Valley	-0.42	-0.64	-0.44	-0.24	0.07	-0.14	-0.84	-0.33	-0.25
Other	0.04	-0.11	-0.09	0.06	0.16	0.11	0.66	0.35	0.01

Table 8. Distribution of Disadvantaged Populations by Region, California, 2010

	Quartile of HPI Score					Percent
Region	Least 1	2	3	Most 4	Sum	Disad-vantaged
Bay Area	3,731,396	1,891,315	1,070,182	395,970	7,088,863	6
Inland Valley	269,521	969,298	1,321,923	1,611,889	4,172,631	39
Los Angeles Co.	2,754,050	2,842,370	3,526,661	3,587,956	12,711,037	28
Sacramento Area	684,681	1,191,661	1,005,423	516,547	3,398,312	15
San Diego	537,842	736,755	608,362	401,649	2,284,608	18
San Joaquin Valley	834,595	959,450	789,767	631,014	3,214,826	20
Other	138,956	705,116	1,065,965	1,976,661	3,886,698	51
Sum	8,951,041	9,295,965	9,388,283	9,121,686	36,756,975	25

All California counties except Alpine had an HPI-eligible census tract. Fifteen counties, mostly in the northern and central Sierras and the Bay Area, did not have any census tracts in the most disadvantaged quartile (Alpine, Amador, Colusa, El Dorado, Inyo, Marin, Mariposa, Modoc, Mono, Napa, Nevada, San Benito, San Mateo, Sonoma, Tuolumne). However, rural areas had a higher proportion of the census tracts in the most disadvantaged quartile (28.6%; 212/742) than urban areas (24.6%, 1736/7049).

Comparison with Other Indices of Disadvantage

The concordance of the most disadvantaged 25% HPI census tracts with those of the most disadvantaged 25% census tracts of CalEnviroScreen, the Hardship Index, the Human Development Index are presented in Table 9 and 10 along with comparisons census tracts below 200% of the federal poverty level and 80% ($49,454) of the median household income.

Table 9. Census Tract Agreement Between HPI and Alternative Indexes

A. CES 3.0

25% Most Disadvantaged

		Y	N	Sum	Sensitivity	Specificity	PA	PPV
HPI	Y	1299	649	1948	0.67	0.89	0.83	0.67
25% Most	N	651	5192	5843				
Disadvantaged	Sum	1950	5841	7791				

B. Hardship Index

25% Most Disadvantaged

		Y	N	Sum	Sensitivity	Specificity	PA	PPV
HPI	Y	1569	379	1948	0.81	0.94	0.90	0.81
25% Most	N	380	5465	5845				
Disadvantaged	Sum	1949	5844	7793				

C. 200% of Federal Poverty Level

		Y	N	Sum	Sensitivity	Specificity	PA	PPV
HPI	Y	1641	307	1948	0.84	0.95	0.92	0.84
25% Most	N	308	5537	5845				
Disadvantaged	Sum	1949	5844	7793				

D. 80% of Median Household Income

		<80%	≥80	Sum	Sensitivity	Specificity	PA	PPV
HPI	Y	1790	158	1948	0.68	0.97	0.87	0.92
25% Most	N	843	5002	5845				
Disadvantaged	Sum	2633	5160	7793				

Note: Colors correspond to map legends in Appendix D
CES, CalEnviroScreen 2.0, HDI, Health Disadvantage Index 1.1; MHI, Median Household Income; PA, Proportion of agreement; PPV, positive predictive value

Table 10. Residential Population in Census Tracts by Agreement Status for HPI and Alternative Indexes

A. CES 3.0

		25% Most Disadvantaged						
HPI		Y	N	Sum	Sensitivity	Specificity	PA	PPV
25% Most Disadvantaged	Y	6,077,925	3,043,761	9,121,686	0.66	0.89	0.83	0.67
	N	3,170,155	24,461,622	27,631,777				
	Sum	9,248,080	27,505,383	36,753,463				

B. Hardship Index

		25% Most Disadvantaged						
		Y	N	Sum	Sensitivity	Specificity	PA	PPV
HPI	Y	7,468,805	1,652,881	9,121,686	0.80	0.94	0.90	0.82
25% Most Disadvantaged	N	1,885,044	25,750,245	27,635,289				
	Sum	9,353,849	27,403,126	36,756,975				

C. 200% of Federal Poverty Level

		Y	N	Sum	Sensitivity	Specificity	PA	PPV
HPI	Y	7,723,774	1,397,912	9,121,686	0.84	0.95	0.92	0.85
25% Most Disadvantaged	N	1,463,802	26,171,487	27,635,289				
	Sum	9,187,576	27,569,399	36,756,975				

D. 80% of Median Household Income

		<80%	≥80	Sum	Sensitivity	Specificity	PA	PPV
HPI	Y	8,325,959	795,727	9,121,686	0.69	0.97	0.88	0.91
25% Most Disadvantaged	N	3,729,531	23,905,758	27,635,289				
	Sum	12,055,490	24,701,485	36,756,975				

Note: Colors correspond to map legends in Appendix D.
CES, CalEnviroScreen 2.0, HDI, Health Disadvantage Index 1.1; MHI, Median Household Income; PA, Proportion of agreement; PPV, positive predictive value

Approximately 650 census tracts, accounting for 3 million Californians, were in disagreement. CES had more California counties than HPI without any census tracts in the quartile of greatest disadvantage (29 vs. 15): Alpine, Amador, Calaveras, Colusa, Del Norte, El Dorado, Glenn, Humboldt, Inyo, Lake, Lassen, Marin, Mariposa, Mendocino, Modoc, Mono, Napa, Nevada, Placer, Plumas, San Benito San Luis Obispo, Santa Barbara, Shasta, Sierra, Siskiyou, Tehama, Trinity, Tuolumne. Sixteen counties had at least one census tract in the most disadvantaged quartile of HPI, but no census tract in the most disadvantaged CES 3.0 quartile. These 16 counties are rural and are from California's north and central coast and northern Sierras:

Calaveras, Del Norte, Glenn, Humboldt, Lake, Lassen, Mendocino, Placer, Plumas, San Luis Obispo, Santa Barbara, Shasta, Sierra, Siskiyou, Tehama, and Trinity.

The positive predictive value of HPI with the indices that emphasized economic status (hardship, poverty, and 80% median household income) ranged from 0.81 to 0.92. Federal poverty level (<200%) had the fewest number of discordant census tracts and population.

Comparison with HDI1.1

HPI and HDI1.1 differ in several ways: the number and types of domains, the methodology for weighting domains, the choice of indicators within domains, the use of more recent, but overlapping data (generally 2011-2015 vs 2008-2012), and the positive framing of the indicators. Of the original 27 HDI1.1 indicators, 10 with an equivalent definition were retained and updated with more recent data: crowded, income, noauto, nowork, poverty, notinhischol, notinpreschool, pm25, renters, and singleparant. Three indicators were identical in definition (parks, retail, treecanopy) but not updated because more recent data were not available. Five indicators had revised definitions and updated data. Educational attainment at age 25 years used a bachelor's degree cut-off in HPI vs. high school in HDI 1.1. Health insurance restricted the population to 18 to 64 year olds in HPI, but was inclusive to all ages in HDI1.1. Supermarket access in HPI used a ¼ mile threshold (vs. 1 mile) for urban areas and retained the 1 mile threshold for rural settings. Housing cost burden was redefined to focus on low-income populations and a higher level of burden (50% vs. 30% of income) in order to be consistent with definitions from the California Department of Housing and Development. The definition of inadequate housing facilities, which only considered lack of kitchen facilities in HDI1.1, also included lack of indoor plumbing in HPI.

Eight HDI 1.1 indicators were discontinued in the HPI, including 5 health outcomes (low birth weight, asthma ER rate, pedestrian injuries, LEB, and disability) and the entire health domain. The 3 other discontinued indicators were traffic density, voting in non-presidential election years, and linguistic isolation.

Although it was not possible to directly assess the independent contributions of updated data and changes in domains and indicators between HDI 1.1 and HPI, it was possible to compare the original HDI 1.1 with a version with the same indicators and domains, but with updated data. This version, called "HDI 2.0", can be used to isolate the impacts of just updating data, and then this version can be compared to HPI to assess the impacts of just changes in methodology. The magnitude of changes can then be compared to the overall changes between HDI 1.1 and HPI (Table 11). In comparisons of the 25% most disadvantaged census tracts of HDI 1.1, HDI 2.0, and HPI, there is a progressive decline in sensitivity as changes progress through updating and methods. The steepest decline in sensitivity with HDI1.1 appears at the stage of updating data (0.85) for HDI 2.0, but change in methods adds to a further decline from HDI 2.0 to HPI (0.79).

Table 11. Comparisons of the Most Disadvantaged 25% Census Tracts in HDI 1.1, HDI1.1 updated ("HDI 2.0"), HPI, and HPI_{Race+}

A. HDI 1.1

		25% Most Disadvantaged						
HDI 2.0		Y	N	Sum	Sensitivity	Specificity	PA	PPV
25% Most	Y	1660	289	1949	0.85	0.95	0.93	0.85
Disadvantaged	N	289	5555	5844				
	Sum	1949	5844	7793				

B. HDI 2.0

		25% Most Disadvantaged						
HPI		Y	N	Sum	Sensitivity	Specificity	PA	PPV
25% Most	Y	1646	302	1948	0.84	0.95	0.92	0.84
Disadvantaged	N	303	5542	5845				
	Sum	1949	5844	7793				

C. HDI 1.1

		25% Most Disadvantaged						
HPI		Y	N	Sum	Sensitivity	Specificity	PA	PPV
25% Most	Y	1539	410	1949	0.79	0.93	0.89	0.79
Disadvantaged	N	409	5435	5844				
	Sum	1948	5845	7793				

D. HPI Race+

		25% Most Disadvantaged						
HPI		Y	N	Sum	Sensitivity	Specificity	PA	PPV
25% Most	Y	1703	245	1948	0.87	0.96	0.94	0.87
Disadvantaged	N	245	5600	5845				
	Sum	1948	5845	7793				

Of the 25% most disadvantaged census (N=1949), 1449 were common to all three versions (HDI 1.1, HDI 2.0, and HPI).

Compared to HDI 1.1, HPI had a higher correlation with LEB (r, 0.56 vs. 0.41) and greater variance-explained (R^2, 31.0% vs. 17.1%).

Race/Ethnicity

Using the same algorithm based on optimizing LEB, domain weights were recalculated for a version of HPI with a race/ethnicity domain (Table 12). In this version of HPI, race/ethnicity (based on the index of dissimilarity) makes important independent contribution (weight 0.13) to the overall HPI score and appears to act across the other social determinants of health.

Table 12. Weighted Quantile Sums Domain Weights

	Weights	
Domain	HPI Race+	HPI
Economic	0.26	0.32
Education	0.16	0.19
Healthcare Access	0.05	0.05
Housing	0.05	0.05
Neighborhood	0.07	0.08
Clean Environment	0.05	0.05
Race/Ethnicity	0.13	—
Social	0.09	0.10
Transportation	0.13	0.16

That is, in the race+ version, the weights for Economic, Education, Neighborhoods, Social, and Transportation domains were attenuated a few percent each compared to the standard HPI without race. Housing, healthcare, and clean environment domains remained at their floors of ~5%. The correlation with LEB and the R^2 value ($r = 0.58$, $R^2_{adj.} = 0.33$) was slightly greater than that of the version without race. Adding the race domain primarily acts to partition the variance among the domains rather than increase the predictive power of the HPI score and life expectancy.

Decision-Support Indicators and Domains

Table 13 presents 44 decision support indicators and their definitions organized into the following domains: health outcomes and health risk behaviors (N=23), climate threats (N=3), Built Environment and Climate Resilience (N=6), social and climate vulnerability (N=9), and other indices of disadvantage (N=4).

Table 13. Variables and Definition, Healthy Places Index (HPI2.0), Decision Support Indicators

Variable Name	Short Name	Definition	Data Source, Year
Health Outcomes			
ARTHRITIS	Arthritis	Crude prevalence of arthritis, adults aged ≥18 Years	CDC[30], 2014
BINGE	Binge Drinking	Crude prevalence of binge drinking, adults aged ≥18 Years	CDC[30], 2014
BPHIGH	High Blood Pressure	Crude prevalence of high blood pressure, adults aged ≥18 Years	CDC[30], 2014
CANCER	Cancer (except skin)	Crude prevalence of cancer (excl. skin cancer), adults aged ≥18 Years	CDC[30], 2014
CASTHMA	Current Asthma	Crude prevalence of current asthma, adults aged >=18 Years	CDC[30], 2014
CHD	Coronary Heart Disease	Crude prevalence of coronary heart disease, adults aged ≥18 Years	CDC[30], 2014
COPD	Chronic Obstructive Pulmonary Disease	Crude prevalence of chronic obstructive pulmonary disease, adults aged ≥18 Years	CDC[30], 2014
CSMOKING	Current Smoking	Crude prevalence of current smoking, adults aged ≥18 Years	CDC[30], 2014
DIABETES	Diabetes	Crude prevalence of diagnosed diabetes, adults aged ≥18 Years	CDC[30], 2014
KIDNEY	Chronic Kidney Disease	Crude prevalence of chronic kidney disease, adults aged ≥18 Years	CDC[30], 2014
LPA	Physical Activity	Crude prevalence of no leisure-time physical activity, adults aged ≥18 yrs	CDC[30], 2014
MHLTH	Mental Health	Crude prevalence of mental health not good for ≥14 days, adults aged ≥18 Years	CDC[30], 2014
OBESITY	Obesity	Crude prevalence of obesity, adults aged ≥18 Years	CDC[30], 2014
PHLTH	Physical Health	Crude prevalence of physical health not good for ≥14 days, adults aged ≥18 Years	CDC[30], 2014
STROKE	Stroke	Crude prevalence of stroke, adults aged ≥18 Years	CDC[30], 2014
pedhurt	Pedestrian traffic injuries	5-year annual average rate of severe and fatal pedestrian injuries per 100,000 population	SWITRS[50], 2006-2010
lbw	Low birth weight	Prevalence of low birth weight infants	CalEPA[48], 2006-2012
asthmaer	Asthma ER rate	Spatially modeled, age-adjusted rate of emergency department (ED) visits for asthma per 10,000	CalEPA[48], 2011-2013
heartattack	Heart Attack ER rate	Spatially modeled, age-adjusted rate of emergency department visits for AMI per 10,000	CalEPA[48], 2011-2013
leb	Life Expectancy	Life expectancy at birth in 2010	VCU, 2010
disabledmental	Cognitive disabilities	Percent of population aged 5 years and older with a cognitive disability	ACS[42], 2011-2015
disabledphysical	Physical disabilities	Percent of population aged 5 years and older with a mental disability	ACS[42], 2011-2015
disabled	Disability	Percent of population with a disability	ACS[42], 2011-2015

Variable Name	Short Name	Definition	Data Source, Year
Climate Threat			
extremeheat	Extreme Heat Days	Projected annual number of extreme heat days at 2070	Cal-Adapt [51], 2017
wildfire	Wildfire risk	Percent of population currently living in very high wildfire risk areas	CalFIRE [52], 2007
sealevel	Sea Level Rise	Percent of population living in sea level rise inundation areas	PacInst [53], 2009
Built Environment and Climate Resilience			
impervsurf	Impervious Surfaces	Percent impervious surface cover	NLCD[44], 2011
uhii	Urban Heat Island	Urban heat island index: sum of 182 day temp. differences (degree-hr) between urban and rural reference	CalEPA [54], 2010
transitaccess	Public Transit Access	Percent of population residing within ½ mile of a major transit stop	CDPH [55], 2012
crime	Violent Crime Rate	Number of Violent Crimes per 1,000 Population	UCR [56], 2013
aircon	Air Conditioning	Percent of households with air conditioning	RAS [57], 2009
traffic	Traffic Density	Sum of traffic volumes adjusted by road segment length (vehicle-kilometers per hour) divided by total road length (kilometers) within 150 meters of the census tract boundary (2013)	CalEPA[48], 2013
Social and Climate Vulnerability			
children	Children under 5 yrs	Percent of population under 5 years of age	ACS[42], 2011-2015
elders	Adults 65 years and older	Percent of population 65 years of age and older	ACS[42], 2011-2015
outdoors	Outdoors workers	Percent of population employed and aged > 16 working outdoors	ACS[42], 2011-2015
(several)	Race/ethnicity	Percent of population by major race/ethnicity group (White, Latino, Black, Asian, Native Am., Native Haw./Pacific Islander, Multiple Race, Other)	Census [58], 2010
englishspeak	English speaking Household	Percentage of households where at least one person 14 years and older speaks English very well	ACS[42], 2011-2015
foreignborn	Foreign Born	Percent of the population born outside the US or US territory	ACS[42], 2011-2015
Gini (city & county)	Income equality	Gini (0 = equality, 1 = income inequality) at county and city level	ACS[42], 2011-2015
voting16	Voter participation	Percent of registered voters voting in 2016 election	UC Berkeley, 2017
iod	Index of Dissimilarity	Index of Dissimilarity County-level African-American Evenness Aggregating Census Blocks	Census [58], 2010

Variable Name	Short Name	Definition	Data Source, Year
Other Indices of Disadvantage			
hi	Hardship Index	Hardship Index	ACS[42], 2011-2015
ces3	CES 3.0 Score	CalEnviroScreen 3.0 Score	CalEPA[48], 2018
mhi80	80% of Median Household Income	<80% of Median Household Income (Yes, No)	ACS[42], 2011-2015

† ACS, American Community Survey; CHAS, Comprehensive Housing Assessment System; CalEPA; California Environmental Protection Agency; CalFIRE, California Department of Forestry and Fire Protection; CDC, Centers for Disease Control; Census, 2010 U.S. Decennial Census; NLCD, National Land Cover Database; RAS, Report on Appliance Saturation; SWITRS, Statewide Integrated Traffic Records System; USDA FARA, U.S. Department of Agriculture Food Access Research Atlas; USEPA, U.S. Environmental Protection Agency; UC Berkeley, Statewide Database; UCR, Uniform Crime Report; VCU, Virginia Commonwealth University

Mapping Application

The HPI website underwent significant expansion in content and functionality. The HPI score, domains, and individual indicators are presented as interactive maps that provide the values and percentile rankings for 1) all 25 HPI indicators, 8 domains and the overall HPI score and 2) 45 decision support indicators (Figure 3).

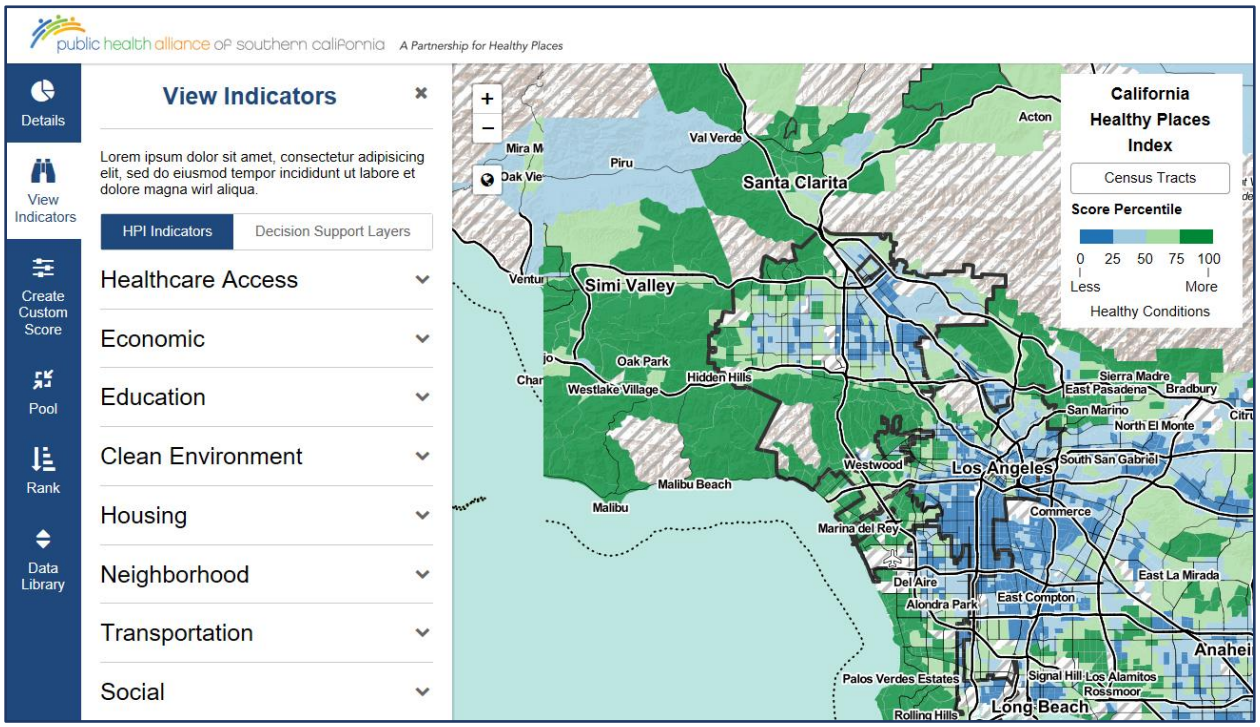


Figure 3. Interactive HPI Map (http://map.healthyplacesindex.org)

A navigation panel allows users to explore potential policy options for improving health by linking individual domains and indicators to briefs that describe 34 policies and 241 policy actions across the 8 HPI domains. In addition to the 25 HPI indicators, the mapping application also provides 45 selectable decision-support data layers covering health outcomes, behavioral risk factors, climate change threats and vulnerabilities, and additional information. The climate change indicators are the same as those used in the CDPH Climate, Health and Equity Program. The mapping application also allows users to pool adjacent census tracts and calculate population-weighted average HPI scores and aggregate census tract data to city, county, and other large geographies. Other features let users create and map their own composite index of individual data layers, upload their own geographies and data, and generate their own community profile report.

Policy Platform

The HPI is built on the premise that improving health outcomes and health equity in California requires both accurate data to understand the conditions that shape health, *and* strong action to shift resources and reshape the places we live. While primarily a data tool, the HPI is designed to facilitate efforts to improve community conditions on the ground. Like the original Health

Disadvantage Index, the HPI tool can be easily used by both public and private organizations to target resources, programs, and other actions to the places where they will have the greatest impact.

New for the HPI, is the addition of a Policy Guide, which links indicator scores to concrete actions that local jurisdictions can use to improve HPI indicators and the underlying community conditions that determine health. For each HPI indicator, the policy guide offers a menu of best practices and emerging policy options that target that indicator. In some cases, these are very directly connected, for instance, policies designed to boost educational opportunity are tied to the education policy action area. Where appropriate, we have also included policies intended to address the root drivers of indicator values, such as policies to improve economic opportunity, which will in turn shape housing affordability, insurance access, and the possibilities open to single-parent households.

Methods

Each policy guide is based on a review of the existing literature connecting a given indicator to health combined with a scan of best practices and conversations with and review from experts in each field (See examples [59-62]). Where possible we have included policies that are evidence based, or are considered industry best-practices. However, since many policy areas in California are rapidly evolving as innovative practices are introduced on the ground, we have also included emerging practices with the potential to improve health. The Policy Guide will continue to evolve along with best practices, and we encourage user feedback and suggestions about policies and resources to include. It should also be noted that we have prioritized policies that specifically address equity and have the potential to close racial, ethnic, gender, economic and geographic disparities in health outcomes.

Use

Each HPI indicator is tied to a variety of policy options that are designed to directly address that indicator, or to influence the root drivers of that indicator. These options are grouped by the general pathway through which they influence health, for instance “Economic Opportunity”, “Health Coverage”, “Transportation”, and then further divided into specific types of intervention, such as “Support Walking and Biking”, or “Plan for Green Communities”. Users can access these options from within the HPI web tool by clicking on a given indicator, or through the stand-along Policy Guide site.

Since many decisions that shape health are made at the local level, and by actors outside of public health, the policy guide is directed to local jurisdictions outside the public health field. Each local jurisdiction is different, and will need different approaches to address long-standing health inequities. The policy guide therefore provides a menu of potential policies for jurisdictions to consider—not a one-size fits all template.

Users wishing to improve community conditions, and their corresponding HPI scores, can use these menus of policies to select a set of policy interventions for further consideration. For each policy, the Policy Guide includes links to guidance documents, examples, and in some cases funding sources. Where possible we have utilized documents produced by government agencies or other authoritative sources, although there also many examples of community-led or smaller scale interventions.

COMMUNICATIONS

In an effort to be responsive to both community and Steering Committee feedback, and to create a tool that resonates with potential users, the Alliance engaged in a re-brand which moved the index away from a frame of "disadvantage" to one of "opportunity." The scale of this shift – impacting not only messaging, but the very name of the index and the way in which data was reported – led the Alliance to seek the counsel of an experienced group of public health communicators in California. This strategic advisory group with 16 members (Appendix A) provided the Alliance with strategic advice which guided a thoughtful re-frame.

The Alliance and VCU worked together with periodic feedback from the communications advisory group and potential end-users to first complete a communications plan establishing a new name, and outlining key messages, target audiences, and strategies for outreach. This document served as guidance for the development of five audience-specific issue briefs, targeting health care providers, California state government, local government officials, the business and financial sector, and community-based organizations and activists.

Major changes included a name shift – from the Health Disadvantage Index to the Healthy Places Index – re-framing reported data to reflect "opportunity" rather than "disadvantage" (for example, highlighting the percentage of the population with a Bachelor's degree or higher, rather than the population without a college degree). Key messages emphasized re-introducing the tool, maintained a focus on the opportunity frame and the utility of the tool for particular audiences, and introduced the concept of social determinants of health to audiences who may not have been already familiar. Examples of messaging below:

- The Healthy Places Index is an important tool to identify places that have the conditions to promote good health. The tool helps us examine how conditions vary by community to help decision-makers target funding, set policy, and pursue other efforts to maximize our wellbeing. The goal is to ensure that all Californians—not just those living in certain communities—have the opportunity to thrive and be healthy.

- The Healthy Places Index is an interactive online data and mapping tool. It allows users to easily understand the conditions that shape health in each neighborhood in California, and to identify the conditions responsible for health inequities, and opportunities to improve outcomes.

- The health of our communities is shaped less by health care than by the conditions in which we live. To have the opportunity to thrive and be healthy, we need access to the building blocks of good health -- good jobs and economic opportunity, quality education, healthy and affordable housing and transportation, clean environments, social supports, health care, and safe neighborhoods that offer green space and support healthy choices.

The five audience-specific briefs are included in Appendix E for reference.

DISCUSSION

Index Development

Some of the key lessons learned in the course of index development were:

- The hybrid approach did achieve its goals: it produced a more predictive index (increased correlation and R^2 with LEB) than the methods used in HDI 1.1 while retaining actionability and understandability by policy makers and other users. This was achieved despite the presence of health outcomes in HDI 1.1.

- Strategies for indicator selection and predictive modeling (WQS) need to take into account indicators with contrary associations with LEB. It may be possible to create a highly predictive index by including indicators and domains with contrary associations, but the coherence of individual indicators are central to actionability and depend not only on the literature but empirical associations in the specific data set.

- Confounding and other possible explanations for contrary associations should be anticipated and explored when using empirical relationships in the dataset of interest to establish domain weights using WQS.

- There is a paucity of indicators available at census tract for clinical and health care quality. Something as basic as physician: population ratio was largely missing. Although the 500 Cities project has indicators of clinical preventive services, its geographic incompleteness leaves a significant data gap.

Communications

While the consensus from the Alliance, VCU, and our advisors was overwhelmingly in support of framing that moved toward “opportunity” and away from “disadvantage” – there were challenges in organizing and presenting some data. Certain variables lend themselves well to a disadvantage orientation, and, when modified, are conceptually abstract and potentially confusing to an end-user.

Likewise, particular end-users, for example, state policymakers, may find a tool oriented to identifying areas of disadvantage more useful for their work. In order to balance their needs with sensitive language and framing, we were deliberate to point out that those areas that were previously thought of as areas of “disadvantage” were better understood as areas of “opportunity” – where lessons learned from areas experiencing more favorable outcome could be applied to improve outcomes in other areas.

There is, however, some inherent confusion around labeling areas in terms of “opportunity.” Those areas that enjoy favorable outcomes, for example, also tend to enjoy ample opportunity for living in health-supportive conditions. At the same time, those areas with less favorable outcomes present an “opportunity” to policymakers, business leaders, and advocates, to strategically support the development of health supportive resources that can work to improve outcomes in an area. We worked to establish very deliberate messaging around reported scores and outcomes, potential uses of the tool, and how communities should be identified and described.

ACKNOWLEDGEMENTS

The California Endowment and Kaiser Permanente are gratefully acknowledged for financial support of this work. Assistance was also provided by the Office of Health Equity, California Department of Public Health.

REFERENCES

1. McLennan D, Barnes H, Noble M, Davies J, Garratt E. *The English Indexes of Deprivation 2010*. London, UK: Department for Communities and Local Government; 2011.https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6320/1870718.pdf
2. Messer LC, Laraia BA, Kaufman JS, Eyster J, Holzman C, J C. The development of a standardized neighborhood deprivation index. *Journal of Urban Health.* 2006;83(6):1041–1062.
3. Singh GK, Azuine RE, Siahpush M, Kogan MD. All-cause and cause-specific mortality among US youth: socioeconomic and rural-urban disparities and international patterns. . *J Urban Health.* 2013;90(3):388-405.
4. Lantz PM, Pritchard A. Socioeconomic indicators that matter for population health. *Prev Chronic Dis.* 2010;7(4):A74.
5. Evans GW, English K. The environment of poverty: multiple stressor exposure, psychophysiological stress, and socioemotional adjustment. *Child Dev.* 2002;73:1238-1248.
6. Ross CE, Mirowsky J. Neighborhood disadvantage, disorder, and health. *Journal of Health and Social Behavior.* 2001;4:258-276.
7. Woolf S, Laudan A, eds. *U.S. Health in International Perspective: Shorter Lives, Poorer Health*. Washington, DC: National Research Council and Institute of Medicine; 2013.http://www.nap.edu/catalog/13497/us-health-in-international-perspective-shorter-lives-poorer-health
8. Diez Rouz AV, Mair C. Neighborhoods and health. *Annals of the New York Academy of Sciences.* 2010;1186:125-145.
9. Kreiger N, Zierler S, Hogan JW, Waterman P, Chen J, Lemieux K, et al. *Geocoding and Measurement of Neighborhood Socioeconomic Position. A U.S. Perspective. In: Kawachi I & Berkman L. Neighborhoods and Health*. Oxford, UK: Oxford University Press; 2003.
10. Townsend P. Deprivation. *Journal of Social Policy.* 1987;16(2):125-146.
11. Pink B. *Socio-Economic Indexes for Areas. Technical Paper 2033.0.55.001*. Sydney, Australia: Australian Bureau of Statistics; 2011.http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/22CEDA8038AF7A0DCA257B3B00116E34/$File/2033.0.55.001%20seifa%202011%20technical%20paper.pdf
12. World Health Organization. *Health Impact Assessment: Determinants of Health*. Geneva: World health Organization; 2015.http://www.who.int/hia/evidence/doh/en/
13. Bhatia R, Maizlish N. *Methodology for a Health Disadvantage Index for California. Final Report (1/21/2016)*. Berkeley, CA and Los Angeles, CA: The Civic Engine and the Public Health Alliance of Southern California; 2016.http://phasocal.org/wp-content/uploads/2016/02/HDI1.1FinalReport2016-01-26.pdf
14. Maizlish N. *Technical Documentation: Health Disadvantage Index (HDI 1.1, 1/21/2016)*. Los Angeles, CA: Public Health Alliance of Southern California; 2016.
15. Centers for Disease Control and Prevention. *Social Determinants of Health: Know What Affects Health*. Atlanta, GA: Centers for Disease Control and Prevention; 2017.https://www.cdc.gov/socialdeterminants/

16. University of Wisconsin Population Health Institute and Robert Wood Johnson Foundation. *County Health Rankings: Our Approach. http://www.countyhealthrankings.org/our-approach;* 2017.
17. U.S. Census Bureau. *2010 Census Geography Note 5, California*. Washington, DC: U.S. Bureau of Census; 2013.http://www2.census.gov/geo/pdfs/reference/Geography_Notes.pdf
18. Booske B, Athens J, Kindig D, Park H, Remington P. *Different Perspectives For Assigning Weights To Determinants of Health*. Madison, WI: University of Wisconsin Population Health Institute; 2010.http://www.countyhealthrankings.org/sites/default/files/differentPerspectivesForAssigningWeightsToDeterminantsOfHealth.pdf
19. Gennings C, Sabo R, Carney E. Identifying subsets of complex mixtures most associated with complex diseases: Polychlorinated biphenyls and endometriosis as a case study. *Epidemiology* 2010;21:S77–S84.
20. Ratcliffe M, Burd C, Holder K, Fields A. *Defining Rural at the U.S. Census Bureau: American Community Survey and Geography Brief ACSGEO-1*. Washington, DC: U.S. Census Bureau; 2016.https://www2.census.gov/geo/pdfs/reference/ua/Defining_Rural.pdf
21. Iceland J, Weinberg D, Steinmetz E. *Racial and Ethnic Residential Segregation in the United States: 1980-2000. CENSR-3*. Washington, DC: U.S. Census Bureau; 2002.https://www.census.gov/prod/2002pubs/censr-3.pdf
22. Massey D, Denton N. The dimensions of racial segregation. *Social Forces.* 1988;67:281-315.
23. Bobo L, Charles C, Krystan M, Simmons A. *The Real Record on Racial Attitudes. In: Marsden, PV, ed. Social Trends in American Life: Findings from the General Social Survey since 1972*. Princeton, NJ: Princeton University Press; 2012.http://scholar.harvard.edu/files/bobo/files/2012_real_record_on_racial_attitudes_social_trends_in_american_life_0.pdf
24. The Lewis Medford Center. *Ethnic Diversity Grows, Neighborhood Integration Lags Behind*. Albany, NY: State University of New York; 2001.http://mumford.albany.edu/census/
25. Glaeser E, Vigdor J. *Racial Segregation in the 2000 Census: Promising News*. Washington, DC: The Brooking Institution Center on Urban & Metropolitan Policy; 2001.https://www.brookings.edu/research/racial-segregation-in-the-2000-census-promising-news/
26. Quinn L, Pawasarat J. *Racial Integration in Urban America: A Block Level Analysis of African American and White Housing Patterns*. Milwaukee, WI: Employment and Training Institute, School of Continuing Education, University of Wisconsin; 2003.https://www4.uwm.edu/eti/integration/integration.pdf
27. Logan J, Stowell J, Oakley D. *Choosing Segregation: Racial Imbalance in American Public Schools, 1990-2000*. Albany, NY: The Lewis Mumford Center, State University of New York; 2002.https://s4.ad.brown.edu/Projects/usschools/reports/report4.pdf
28. U.S. Census Bureau. *2010 Census Redistricting Data (Public Law 94-171) Summary File — California [machine-readable data files]*. Washington, DC: U.S. Census Bureau; 2011.https://www.census.gov/prod/cen2010/doc/pl94-171.pdf
29. Zhang X, Holt J, Yun S, Lu H, Greenlund K, Croft J. Validation of multilevel regression and poststratification methodology for small area estimation of health indicators from the Behavioral Risk Factor Surveillance System. . *Am J Epidemiol.* 2015;182:127–137.
30. Centers for Disease Control and Prevention. *500 Cities: Local Data for Better Health* Atlanta, GA; 2017.https://www.cdc.gov/500cities/
31. Faust J, August L, Alexeeff G, Bangia K, Cendak R, Cheung-Sutton E, et al. *California Communities Environmental Health Screening Tool, Version 3.0 (June 2018 update): Guidance and Screening Tool*. Sacramento, CA: Office of Environmental Health Hazard

Assessment, Californa Environmental Protection Agency; 2014.http://oehha.ca.gov/ej/pdf/CES20FinalReportUpdateOct2014.pdf
32. Nathan RP, Adams CF. Four Perspectives on Urban Hardship. *Political Science Quarterly.* 1989;104(3):483-508.
33. Wright DJ, Montiel LM. *Divided They Fall: Hardship in America's Cities and Suburbs*. Albany, NY: The Nelson A. Rockefeller Institute of Government; 2007.
34. Los Angeles County Department of Public Health. *Comunity Health Assessment*. Los Angeles: Office of Planning, Evaluation, and Development, Los Angeles County Department of Public Health; 2015.http://www.publichealth.lacounty.gov/plan/docs/CHA_CHIP/LACDPHCommunityHealthAssessment2015.pdf
35. Office of the Assistant Secretary for Planning and Evaluation. *2015 Poverty Guidelines*. Washington, DC: U.S. Department of Health and Human Services; 2015.
36. Center for Family Health Women, Infants and Children Program. *State, County and Regional-Level Estimates of WIC Eligibles and Program Reach, California 2011*. Sacramento, CA: California Department of Public Health; 2016.
37. Office of Policy Development & Research. *FY 2016 HUD Income Limits Briefing Material*: U.S. Department of Housing and Urban Development; 2016.https://www.huduser.gov/portal/datasets/il/il16/IncomeLimitsBriefingMaterial-FY16.pdf
38. Bates L. *State Income Limits for 2015*. Sacramento, CA: California Department of Housing and Community Development; 2015.http://www.hcd.ca.gov/housing-policy-development/housing-resource-center/reports/state/inc2k15.pdf
39. State Water Resources Control Board. *California Regulations Related to Drinking Water. §63000.25. Disadvantaged Community*. Sacramento, CA: State Water Resources Control Board; 2015.http://www.waterboards.ca.gov
40. Fenelon A, Chinn J, Anderson R. A comprehensive analysis of the mortality experience of hispanic subgroups in the United States: Variation by age, country of origin, and nativity. *SSM - Population Health* 2017;3:245–254.
41. Hastings K, Jose P, Kapphahn K, Frank A, Goldstein B, Thompson C, et al. Leading causes of death among Asian American subgroups (2003–2011). *PLOS One.* 2015;10:e0124341.
42. U.S. Census Bureau. *American Community Survey*. DS. Washington, DC: U.S. Census Bureau; 2011-2015.http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml
44. Homer CG, Dewitz JA, Yang L, Jin S, Danielson P, Xian G, et al. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing* 2015;81(5):345-354.
48. Faust J, August L, Bangia K, Schmitz R, Galaviz V, Leichty J, et al. *CalEnviroScreen 3.0: Update to the California Communities Environmental Health Screening Tool* Sacramento, CA: Office of Environmental Health Hazard Assessment, California Environmental Protection Agency; 2017.http://oehha.ca.gov/media/downloads/calenviroscreen/report/ces3report.pdf
50. California Highway Patrol. *Statewide Integrated Traffic Records System (SWITRS)*. Sacramento: California Highway Patrol; 2010.www.chp.ca.gov/switrs/index.htm, accessed 8/2/2010
51. California Energy Commission. *Cal-Adapt: Exploring California's Climate Change Research*. Sacranebnto, CA; 2017.http://cal-adapt.org/
52. California Department of Forestry and Fire Protection. *Fire Hazard Severity Zone Map Update Project, State Responsibility Area (SRA) and Local Responsibility Area (LRA)*. Sacramento, CA: California Department of Forestry and Fire Protection; 2007.http://www.fire.ca.gov/fire_prevention/fire_prevention_wildland_statewide.php

53. Heberger M, Cooley H, Herrera P, Gleick P, Moore E. *The Impacts of Sea Level Rise on the California Coast*. Oakland, CA: Pacific Institute; 2009.http://pacinst.org/the-impacts-of-sea-level-rise-on-the-california-coast-gis-data-downloads/
54. Taha H. *Creating and Mapping an Urban Heat Island Index for California. Final Report*. Martinez, CA: Altostratus Inc.; 2015.http://www.calepa.ca.gov/UrbanHeat/Report/Report.pdf
55. Office of Health Equity. *Healthy Communities Data and Indicators Project*. Sacramento, CA: California Department of Public Health; 2015.https://www.cdph.ca.gov/programs/Pages/HealthyCommunityIndicators.aspx
56. Federal Bureau of Investigation. *Uniform Crime Reports*. Washington, DC: Federal Bureau of Investigation; 2013.https://oag.ca.gov/crime/cjsc/criminal-justice-profiles
57. Palmgren C, Stevens N, Goldberg M, Barnes R, Rothkin K. *California Residential Appliance Saturation Survey*. Oakland, CA: KEMA, Inc.; 2009.
58. U.S. Census Bureau. *U.S. Decennial Census*. DS. Washington, DC: U.S. Census Bureau; 2010.http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml
59. United States Interagency Council on Homelessness. *Criminal Justice Reform*. Washington, D.C.: United States Interagency Council on Homelessness; 2017.https://www.usich.gov/solutions/criminal-justice
60. Southern California Association of Governments. *Affordable Housing and Anti-Displacement Policies*. Los Angeles, CA: Southern California Association of Governments.http://www.scag.ca.gov/Documents/ComprehensiveGuideToLocalAffordableHousingPolicy.pdf
61. California Office of Planning and Research. *General Plan Guidelines*. Sacramento, CA: California Office of Planning and Research; 2017.https://www.opr.ca.gov/planning/general-plan/guidelines.html
62. California Department of Parks and Recreation. *Statewide Park Development and Community Revitalization Program*. Sacramento, CA: California Department of Parks and Recreation; 2016.http://www.parks.ca.gov/?page_id=26025

APPENDICES

A. Membership on the HPI Steering Committee and Communications Advisory Group

Members of the HPI Steering Committee
Meileen Acosta, Solano County Department of Public Health
Matt Beyers, Alameda County Health Department
Dulce Bustamante, Office of Health Equity, California Department of Public Health
Charlene Contreras, Los Angeles County Department of Public Health
Dave Dauphine, California Department of Public Health
Satvinder Dhaliwal, Fresno County Department of Public Health
Scott Fujimoto, Center for Health Statistics and Informatics, California Department of Public Health
Elizabeth Gazarek, Kings County Department of Public Health
Solange Gould, Office of Health Equity, California Department of Public Health
Wendy Hetherington, Riverside County Department of Public Health
David Holstius, Bay Area Air Quality Management District
Trav Ichinose, Orange County Department of Public Health
Melissa Jones, Executive Director, Bay Area Regional Health Inequities Initiative
Tammy Lee, Alameda County Health Department
Meredith Milet, Office of Health Equity, California Department of Public Health
Julie Nagasako, Fusion Center for Strategic Development, California Department of Public Health
Racheline Napier, Riverside County Public Health Department
Linda Rudolph, Director, Center for Climate Change and Health, Public Health Institute
Michael Samuel, Fusion Center for Strategic Development, California Department of Public Health
Paul Simon, Los Angeles County Department of Public Health
Latesa Slone, California Department of Public Health
Jason Vargo, Office of Health Equity, California Department of Public Health

Members of the Communications Advisory Group
Elizabeth Baca, California Office of Planning and Research
Rye Baerg, Southern California Council of Governments
Dalila Butler, PolicyLink
Lianne Dillon and Julia Caplan, California Strategic Growth Council, Health In All Policies Team
Solange Gould and Meredith Millet, California Department of Public Health, Office of Health Equity
Melissa Jones, Bay Area Regional Health Inequities Initiative
Jennifer Lopez, Kaiser
Kate Meis and Paul Zykofsky, Local Government Commission
Sarah Reyes, The California Endowment
Kathi Shaff, Berkeley Media Studies Group
Kerri Timmer, Sierra Business Council
Elva Yanez and Rob Baird, Prevention Institute

B. Data Dictionary and Source Data Variable Transformations for HPI 2.0 Files

1. HPI Indicators (Alphabetic Order)

Variable Name	Data Source	Table	Variable(s)
abovepoverty	ACS5YR2015	S1701	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/S1701/0400000US06.14000 HC01_EST_VC56 Total; Estimate; ALL INDIVIDUALS WITH INCOME BELOW THE FOLLOWING POVERTY RATIOS - 200 percent of poverty level HC01_MOE_VC56 Total; Margin of Error; ALL INDIVIDUALS WITH INCOME BELOW THE FOLLOWING POVERTY RATIOS - 200 percent of poverty level HC01_EST_VC01 Total; Estimate; Population for whom poverty status is determined HC01_MOE_VC01 Total; Margin of Error; Population for whom poverty status is determined Numerator = HC01_EST_VC01 – HC01_EST_VC56 Denominator = HC01_EST_VC01 Percent = Numerator/denominator x 100 se(numerator, A)=HC01_MOE_VC56/1.645 se(denominator, B)=HC01_MOE_VC01/1.645 se_pct= Let $P = \frac{A}{B}$. $SE(P) = \frac{1}{B}\sqrt{SE(A)^2 - P^2 \times SE(B)^2}$
alcoffsale	CDPH/ABC		URL: https://www.cdph.ca.gov/Programs/OHE/Pages/Healthy-Communities-Data-and-Indicators-Project-(HCI).aspx HCI_AlcoholOutletsQ_774_CA_RE_CO_CD_PL_CT-A-N-5-16-14.xlsx HCI_AlcoholOutletsQ_774_CO_CD_PL_CT-O-Y-5-16-14.xlsx Excel files must be appended and filtered for race_eth_code==9 & geotype=="CT" & (license_type == "Off_sale") alcoffsale_pct = percent alcoffsale_se_pct = se Note: must contact program; data not available on new website as of 11/10/2017
automobile	ACS5YR2015	DP04	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/DP04/0400000US06.14000

			HC03_VC85 Percent; VEHICLES AVAILABLE - Occupied housing units - No vehicles available HC04_VC85 Percent Margin of Error; VEHICLES AVAILABLE - Occupied housing units - No vehicles available HC01_VC84 Estimate; VEHICLES AVAILABLE - Occupied housing units HC01_VC85 Estimate; VEHICLES AVAILABLE - Occupied housing units - No vehicles available Numerator=100- HC01_VC85 Denominator= HC01_VC84 Percent = 100 – HC03_VC85 se_pct=HC04_VC85/1.645
bachelorsed	ACS5YR2015	DP02	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/DP02/0400000US06.14000 HC03_VC96 Percent; EDUCATIONAL ATTAINMENT - Percent bachelor's degree or higher HC04_VC96 Percent Margin of Error; EDUCATIONAL ATTAINMENT - Percent bachelor's degree or higher HC01_VC85 Estimate; Total: Population 25 years and over Numerator=(Percent/100)*Denominator Denominator=HC01_VC85 Percent=HC03_VC96 se_pct=HC04_VC96/1.645
commute	ACS5YR2015	B08301	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/B08301/0400000US06.14000 HD01_VD01 Estimate; Total: HD02_VD01 Margin of Error; Total: HD01_VD10 Estimate; Public transportation (excluding taxicab) HD02_VD10 Margin of Error; Public transportation (excluding taxicab) HD01_VD18 Estimate; Bicycle HD02_VD18 Margin of Error; Bicycle HD01_VD19 Estimate; Walked HD02_VD19 Margin of Error; Walked HD01_VD21 Estimate; Worked at home HD02_VD21 Margin of Error; Worked at home Numerator=HD01_VD10+HD01_VD18+HD01_VD19 Denominator=HD01_VD01-HD01_VD21

			Percent=Numerator/Denominator*100 se(numerator, A)=sqrt((HD02_VD10/1.645)^2 + (HD02_VD18/1.645)^2 + (HD02_VD19/1.645)^2) se(denominator,B)=SQRT((HD02_VD01/1.645)^2-(HD02_VD21/1.645)^2) se_pct= Let $P = \frac{A}{B}$. $$SE(P) = \frac{1}{B}\sqrt{SE(A)^2 - P^2 \times SE(B)^2}$$
dieselpm	CES 3.0	Excel File	URL: https://oehha.ca.gov/media/downloads/calenviroscreen/document/ces3results.xlsx Dieselpm = Diesel PM (Diesel PM emissions from on-road and non-road sources)
employed	ACS5YR2015	S2301	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/S2301/0400000US06.14000 HC03_EST_VC26 Employment/Population Ratio; Estimate; Population 20 to 64 years HC03_MOE_VC26 Employment/Population Ratio; Margin of Error; Population 20 to 64 years HC01_EST_VC26 Total; Estimate; Population 20 to 64 years Numerator=(Percent/100)*Denominator Denominator=HC01_EST_VC26 Percent=HC03_EST_VC26 se_pct=HC03_MOE_VC26/1.645
h20contam	CES 3.0	Excel File	URL: https://oehha.ca.gov/media/downloads/calenviroscreen/document/ces3results.xlsx H20contam = Drinking Water (Drinking water contaminant index for selected contaminants)
homeownership	ACS5YR2015	DP04	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/DP04/0400000US06.14000 HC03_VC65 Percent; HOUSING TENURE - Occupied housing units - Owner-occupied HC04_VC65 Percent Margin of Error; HOUSING TENURE - Occupied housing units - Owner-occupied Percent = HC03_VC65 se_pct=HC04_VC65/1.645
houserepair	CHAS_2010-2014_Tables15A_15B_15C	Table 15A, Table 15B,	URL: https://www.huduser.gov/portal/datasets/cp.html#2006-2014_data T15A_est3 Owner occupied with mortgage AND has complete kitchen and plumbing facilities T15A_moe3

		Table 15C	T15B_est3 Owner occupied with no mortgage AND has complete kitchen and plumbing facilities T15B_moe3 Margin of Error; Owner occupied with no mortgage AND has complete kitchen and plumbing facilities T15C_est3 Renter occupied AND has complete kitchen and plumbing facilities T15C_moe3 Margin of Error; Renter occupied AND has complete kitchen and plumbing facilities T15A_est1 Owner occupied with mortgage T15A_moe1 Margin of Error; Owner occupied with mortgage T15B_est1 Owner occupied with no mortgage T15B_moe1 Margin of Error; Owner occupied with no mortgage T15C_est1 Renter occupied T15C_moe1 Margin of Error; Renter occupied Numerator = T15A_est3 + T15B_est3 + T15C_est3 Denominator = T15A_est1 +T15B_est1 + T15C_est1 Percent = (Numerator/denominator) x 100 se(numerator, A)=SQRT((T15A_moe3/1.645)^2 + (T15B_moe3/1.645)^2 + (T15C_moe3/1.645)^2) se(denominator, B)=SQRT((T15A_moe1/1.645)^2 + (T15B_moe1/1.645)^2 + (T15C_moe1/1.645)^2) se_pct= Let $P = \frac{A}{B}$. $$SE(P) = \frac{1}{B}\sqrt{SE(A)^2 - P^2 \times SE(B)^2}$$
income	ACS5YR2015	DP03	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/DP03/0400000US06.14000 HC01_VC85 Estimate; INCOME AND BENEFITS (IN 2015 INFLATION-ADJUSTED DOLLARS) - Total households - Median household income (dollars) HC02_VC85 Margin of Error; INCOME AND BENEFITS (IN 2015 INFLATION-ADJUSTED DOLLARS) - Total households - Median household income (dollars) Percent=HC01_VC85 se_pct=HC02_VC85/1.645
inhighschool	ACS5YR2015	S1401	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/S1401/0400000US06.14000

			HC02_EST_VC25 Percent; Estimate; Population 15 to 17 - 15 to 17 year olds enrolled in school HC02_MOE_VC25 Percent; Margin of Error; Population 15 to 17 - 15 to 17 year olds enrolled in school HC01_EST_VC25 Total; Estimate; Population 15 to 17 - 15 to 17 year olds enrolled in school Numerator= HC01_EST_VC25 Denominator=Numerator/Percent*100 Percent=HC02_EST_VC25 se_pct= HC02_MOE_VC25/1.645
inpreschool	ACS5YR2015	S1401	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/S1401/0400000US06.14000 HC02_EST_VC16 Percent; Estimate; Population 3 to 4 years - 3 to 4 year olds enrolled in school HC02_MOE_VC16 Percent; Margin of Error; Population 3 to 4 years - 3 to 4 year olds enrolled in school HC01_EST_VC16 Total; Estimate; Population 3 to 4 years - 3 to 4 year olds enrolled in school Numerator= HC01_EST_VC16 Denominator= Numerator/Percent*100 Percent=HC02_EST_VC16 se_pct=HC02_MOE_V16/1.645
insured	ACS5YR2015	S2701	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/S2701/0400000US06.14000 HC03_EST_VC07 Percent Insured; Estimate; AGE - 18 to 64 years HC03_MOE_VC07 Percent Insured; Margin of Error; AGE - 18 to 64 years HC02_EST_VC07 Insured; Estimate; AGE - 18 to 64 years HC01_EST_VC07 Total; Estimate; AGE - 18 to 64 years Numerator= HC02_EST_VC07 Denominator= HC01_EST_VC07 Percent= HC03_EST_VC07 se_pct= HC03_MOE_VC07/1.645
rentsevere and ownsevere	CHAS, 2010-2014	Table 8	URL: https://www.huduser.gov/portal/datasets/cp.html See table below for variable names Own Severe Numerator = T8_est10 + T8_est23 + T8_est36 Denominator = T8_est2

			Percent = 100*Numerator/denominator SE denominator = T8_moe2/1.645 SE numerator = sqrt[(T8_moe10/1.645)^2 + (T8_moe23/1.645)^2 + (T8_moe36/1.645)^2] SE percent = (1/ T8_est2) * sqrt[(SE$_{num}$)^2 - (Numerator^2/Denominator^2)*(SE$_{denom.}$)^2] <u>Rent Severe</u> Numerator = T8_est76 + T8_est89 + T8_est102 Denominator = T8_est68 Percent = 100*Numerator/denominator SE denominator = T8_moe68/1.645 SE numerator = sqrt[(T8_moe76/1.645)^2 + (T8_moe89/1.645)^2 + (T8_moe102/1.645)^2] SE percent = (1/ T8_est68) * sqrt[(SE$_{num}$)^2 - (Numerator^2/Denominator^2)*(SE$_{denom.}$)^2]

	Owners				Renters			
	Cost Burden				Cost Burden			
Income	>50%	>50%	Total		>50%	>50%	Total	
<80% HAFMI	Estimate	MOE	Estimate	MOE	Estimate	MOE	Estimate	MOE
<30	T8_est10	T8_moe10	T8_est3		T8_est76	T8_moe76	T8_est69	
30-50	T8_est23	T8_moe23	T8_est16		T8_est89	T8_moet89	T8_est82	
50-80	T8_est36	T8_moe36	T8_est29		T8_est102	T8_moe102	T8_est95	
>80% HAFMI								
80-100	T8_est49		T8_est42		T8_est115		T8_est108	
>100	T8_est62		T8_est55		T8_est128		T8_est121	
Total			T8_est2	T8_moe2			T8_est68	T8_moe68

T8_est1 = total occupied housing units

ozone	CES 3.0	Excel File	URL: https://oehha.ca.gov/media/downloads/calenviroscreen/document/ces3results.xlsx ozone = Amount of daily maximum 8 hour Ozone concentration
parkaccess	CDPH		URL: https://www.cdph.ca.gov/Programs/OHE/Pages/Healthy-Communities-Data-and-Indicators-Project-(HCI).aspx Download ParkBeachOpen10_output4-12-13.zip from Filter Excel file for CT and Total Race/ethnicity, copy and save as .csv file Denominator = pop2010 Numerator = pop_park_acc Percent = 1- numerator/denominator se_pct = se

pm25	CES 3.0	Excel File	URL: https://oehha.ca.gov/media/downloads/calenviroscreen/document/ces3results.xlsx Pm25 = PM2.5 (Annual mean PM 2.5 concentrations)
retail	USEPA	SmartLocationDatabase 2.0 (2013)	http://www2.epa.gov/smartgrowth/smart-location-mapping#SLD Note: files is at census block group (CBG) and must be aggregated to census tract D1C8_Ret10 Gross retail (8-tier) employment density (jobs/acre) on unprotected land D1C8_Ent10 Gross entertainment (8-tier) employment density (jobs/acre) on unprotected land D1C8_Ed10 Gross education(8-tier) employment density (jobs/acre) on unprotected land Ac_Unpr Total land area in acres that is not protected from development (i.e., not a park or conservation area) Percent = D1c8_Ret10 + D1c8_Ent10 + D1c8_Ed10 Denominator = Ac_Unpr*Denominator Numerator = percent*
supermkts	USDA Food Access research Atlas, 2015		URL: https://www.ers.usda.gov/data-products/food-access-research-atlas/download-the-data/ lapophalf: Population count beyond 1/2 mile from supermarket (numerator urban) lapop1: Population count beyond 1 mile from supermarket (numerator ruarl) POP2010: Population count from 2010 census (denominator) UrbanType (HDI/Census) urban(urban_area), rural (urban_cluster, rural) lapophalf_pct <- 100*lapophalf/POP2010 for urban_area lapop1pct <- 100*lapop1/POP2010 for urban cluster and rural se_pct = sqrt(percent*(1-percent)/denominator) percentiles only calculated for HDI eligible census tracts
treecanopy	CDPH	BRACE	Download BRACE_TreeCanopy_458_CT_PL_CO_RE_CA__02DEC15.xlsx from: https://www.cdph.ca.gov/programs/Pages/HealthyCommunityIndicators.aspx Filter for race_eth_name (Total), geotype (CT), and strata_level_name (population-weighted), copy and paste as .csv percent = percent se_pct =percent_se rse_pct = percent_rse
twoparent	ACS5YR2015	B09008	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/B09008/0400000US06.14000 HD01_VD03 Estimate; Unmarried partner of householder present: - In family households: HD02_VD03 Margin of Error; Unmarried partner of householder present: - In family households:

			HD01_VD09 Estimate; No unmarried partner of householder present: - In family households: - In married-couple family HD02_VD09 Margin of Error; No unmarried partner of householder present: - In family households: - In married-couple family HD01_VD01 Estimate; Total: HD02_VD01 Margin of Error; Total: Numerator = HD01_VD03 + HD01_VD09 Denominator = HD01_VD01 Percent = Numerator/denominator x 100 se_pct=SQRT((HD02_VD03/1.645)^2+(HD02_VD09/1.645)^2)/100
uncrowded	ACS5YR2015	DP04	URL:https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/DP04/0400000US06.14000 HC03_VC113 Percent; OCCUPANTS PER ROOM - Occupied housing units - 1.00 or less HC04_VC113 Percent Margin of Error; OCCUPANTS PER ROOM - Occupied housing units - 1.00 or less HC01_VC113 Estimate; OCCUPANTS PER ROOM - Occupied housing units - 1.00 or less HC01_VC112 Estimate; OCCUPANTS PER ROOM - Occupied housing units Percent=HC03_VC113 Numerator=HC01_VC113 Denominator=HC01_VC112 se_pct= HC04_VC113/1.645
voting	UCB		URL: http://statewidedatabase.org/d10/g12_geo_conv.html Registration crosswalk file (state_g12_rg_blk_map.csv) was right-joined on RGPREC_KEY to voter file (state_g12_voters_by_g12_rgprec.csv). denominator = BLKREG numerator = TOTREG_R (voters) percent = 1-numerator/denominator se_pct = se_pct = sqrt(percent*(1-percent)/denominator) Note: because the allocation of precinct to block is imperfect some parts of precincts were not allocated to blocks (precinct registration = 17,981,054 vs. block registration = 17,840,280; difference = 140,774, 1%) Method of analysis suggested by Janine Heiser, UCB (510) 624-9086 See technical documentation HDI 1.1, Appendix B for step-by-step procedures

2. Decision Support Indicators (Alphabetic Order)

Variable Name	Data Source	Table	Variable(s)
aircon	HCI2013		URL: https://archive.cdph.ca.gov/programs/Pages/CalBRACE-Indicators2017.aspx Manually filtered BRACE_AirConditioning_797_CO_RE_CA.xlsx for county), added FIPS for to create BRACE_AirConditioningCO.csv, which was reformatted in R (HDI_BRACE_DS_Indicators2017-07-XX.R)
asthmaer	CES 3.0	Excel File	URL: https://oehha.ca.gov/media/downloads/calenviroscreen/document/ces3results.xlsx asthmaer = asthma (Age-adjusted rate of emergency department visits for asthma)
children	ACS_15_5YR	S0101	URL: http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml (download) HC01_EST_VC01 Total; Estimate; Total population HC01_EST_VC03 Total; Estimate; AGE - Under 5 years (percent) HC01_MOE_VC03 Total; Margin of Error; AGE - Under 5 years Denominator = HC01_EST_VC01 Numerator = HC01_EST_VC01* HC01_EST_VC03/100 Percent = HC01_EST_VC03 pct_se = HC01_MOE_VC03/1.645
crime_rate	UCR/CDPH		URL: https://archive.cdph.ca.gov/programs/Pages/CalBRACE-Indicators2017.aspx Manually filtered HCI_CrimePL2013_21OCT15.csv crime_rate = rate crime_se_pct = se
disability	ACS_15_5YR	S1810	URL: http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml (download) Total disability HC01_EST_VC01 Total; Estimate; Total civilian noninstitutionalized population HC02_EST_VC01 With a disability; Estimate; Total civilian noninstitutionalized population HC03_EST_VC01 Percent with a disability; Estimate; Total civilian noninst. population HC03_MOE_VC01 Percent with a disability; Margin of Error; Total civilian noninst.population

			Numerator = HC02_EST_VC01 Denominator = HC01_EST_VC01 Percent = HC03_EST_VC01 se_pct = HC03_MOE_VC01/1.645 Cognitive Disability (5 years and older) *(HC01_EST_VC47 Total; Estimate; DISABILITY TYPE BY DETAILED AGE - With a cognitive difficulty - all values missing)* HC01_EST_VC19 Total; Estimate; AGE - Under 5 years HC02_EST_VC47 With a disability; Estimate; DISABILITY TYPE BY DETAILED AGE - With a cognitive difficulty HC03_EST_VC47 Percent with a disability; Estimate; DISABILITY TYPE BY DETAILED AGE - With a cognitive difficulty HC03_MOE_VC47 Percent with a disability; Margin of Error; DISABILITY TYPE BY DETAILED AGE - With a cognitive difficulty Numerator = HC02_EST_VC47 Denominator = HC01_EST_VC01 - HC01_EST_VC19 Percent = HC03_EST_VC47 se_pct = HC03_MOE_VC47/1.645 *Physical Disability (5 years and older)* *HC01_EST_VC55 Total; Estimate; DISABILITY TYPE BY DETAILED AGE - With an ambulatory difficulty - all missing values)* HC01_EST_VC19 Total; Estimate; AGE - Under 5 years HC02_EST_VC55 With a disability; Estimate; DISABILITY TYPE BY DETAILED AGE - With an ambulatory difficulty HC03_EST_VC55 Percent with a disability; Estimate; DISABILITY TYPE BY DETAILED AGE - With an ambulatory difficulty HC03_MOE_VC55 Percent with a disability; Margin of Error; DISABILITY TYPE BY DETAILED AGE - With an ambulatory difficulty Numerator = HC02_EST_VC55 Denominator = HC01_EST_VC01 - HC01_EST_VC19 Percent = HC03_MOE_VC55 se_pct = HC03_MOE_VC55 /1.645 ACS 2011-2015 file calculated cognitive and ambulatory disability directly (without age stratification, which was done in the 2008-2012 analysis by Jacqui Chan

Variable Name	Data Source	Table	Variable(s)
Elders	ACS_15_5YR	S0101	URL: http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml (download) HC01_EST_VC01 Total; Estimate; Total population HC01_EST_VC31 Total; Estimate; SELECTED AGE CATEGORIES - 65 years and over (percent) HC01_MOE_VC31 Total; Margin of Error; SELECTED AGE CATEGORIES - 65 years and over Denominator = HC01_EST_VC01 Numerator = HC01_EST_VC31* HC01_EST_VC03/100 Percent = HC01_EST_VC31 pct_se = HC01_MOE_VC31/1.645
englishspeak	ACS5YR2015	B16002	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/B16002/0400000US06.14000 HD01_VD04 Estimate; Total: - Spanish: - Limited English speaking household HD02_VD04 Margin of Error; Total: - Spanish: - Limited English speaking household HD01_VD07 Estimate; Total: - Other Indo-European languages: - Limited English speaking household HD02_VD07 Margin of Error; Total: - Other Indo-European languages: - Limited English speaking household HD01_VD10 Estimate; Total: - Asian and Pacific Island languages: - Limited English speaking household HD02_VD10 Margin of Error; Total: - Asian and Pacific Island languages: - Limited English speaking household HD01_VD13 Estimate; Total: - Other languages: - Limited English speaking household HD02_VD13 Margin of Error; Total: - Other languages: - Limited English speaking household HD01_VD01 Estimate; Total: HD02_VD01 Margin of Error; Total: Numerator =Denominator- (HD01_VD04 + HD01_VD07 + HD01_VD10 + HD01_VD13) Denominator = HD01_VD01 Percent =(100- Numerator/denominator) x 100 se(numerator, A)= sqrt((HD02_VD04/1.645)^2 + (HD02_VD07/1.645)^2 + (HD02_VD10/1.645)^2 + (HD02_VD13/1.645)^2) se(denominator, B)=HD02_VD01/1.645 se_pct=

			Let $P = \frac{A}{B}$. $$SE(P) = \frac{1}{B}\sqrt{SE(A)^2 - P^2 \times SE(B)^2}$$
Five Hundred Cities: ARTHRITIS BPHIGH CANCER CASTHMA CHD COPD DIABETES KIDNEY MHLTH PHLTH STROKE LPA CSMOKING BINGE OBESITY	500 Cities (CDC)	November 30, 2016	UR: https://www.cdc.gov/500cities/, https://chronicdata.cdc.gov/500-Cities/500-Cities-Local-Data-for-Better-Health/6vp6-wxuq To create an orthogonal file of unduplicated census tracts for crude prevalence of all the outcomes, the file was filtered in Excel for state (StateAbbrev=CA), census tract (GeographicLevel=Census Tract), and crude prevalence DataValueTypeID= CrdPrev Crude prevalence in adults > 18 years = CrdPrev
foreignborn	ACS5YR2015	B05002	URL: https://factfinder.census.gov/bkmk/table/1.0/en/ACS/15_5YR/B05002/0400000US06.14000 HD01_VD13 Estimate; Foreign born: HD02_VD13 Margin of Error; Foreign born: HD01_VD01 Estimate; Total: HD02_VD01 Margin of Error; Total: Numerator = HD01_VD13 Denominator = HD01_VD01 Percent = Numerator/denominator x 100 se(numerator, A)= HD02_VD13/1.645 se(denominator, B)=HD02_VD01/1.645 se_pct= Let $P = \frac{A}{B}$. $$SE(P) = \frac{1}{B}\sqrt{SE(A)^2 - P^2 \times SE(B)^2}$$
Gini	ACS_15_5YR	B19083	URL: http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml (download)

	(place) (County)		gini_pct = HD01_VD01 gini_se_pct = HD02_VD01/1.645
heartattack			URL: https://oehha.ca.gov/media/downloads/calenviroscreen/document/ces3results.xlsx heartattack = Cardiovascular.Disease
extremeheat	CalAdapt (beta, 6/23/17)		URL: http://beta.cal-adapt.org/tools/extreme-heat/#climatevar=tasmax&scenario=rcp45&lat=37.68014&lng=-121.90565&zoom=9&boundary=counties&units=fahrenheit Manual transcription of county results from HadGEM2-ES model RCP8.5 scenario, averaged for 2070-2099
impervsurf	National Land Cover Database 2011		URL: https://archive.cdph.ca.gov/programs/Pages/CalBRACE-Indicators2017.aspx and http://www.mrlc.gov/nlcd11_data.php Manually filtered BRACE_SLR_784_CT_PL_CO_RE_CA_11-1-2016.xlsx on Total race/ethnicity and census tract geography to create csv file BRACE_Wildfire_CT8-11-2016.csv, which was reformatted in HDI_BRACE_DS_Indicators2017-07-XX.R selected only population-weighted area
lbw	CES 3.0	Excel File	URL: https://oehha.ca.gov/media/downloads/calenviroscreen/document/ces3results.xlsx lbw = Percent low birth weight
LEB	VCU		http://www.societyhealth.vcu.edu/ Data provided by Virginia Commonwealth University Center for Society and health
outdoors	ACS5YR2015	S2401	URL: https://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t#none HC01_EST_VC01 Total; Estimate; Civilian employed population 16 years and over HC01_MOE_VC01 Total; Margin of Error; Civilian employed population 16 years and over HC01_EST_VC30 Total; Estimate; farming, fishing, and forestry occupations HC01_EST_VC31 Total; Estimate; Construction and extraction occupations HC01_MOE_VC30 Total; Margin of Error; Farming, fishing, and forestry occupations HC01_MOE_VC31 Total; Margin of Error; Natural resources, construction, and maintenance occupations: - Construction and extraction occupations Numerator = HC01_EST_VC30 + HC01_EST_VC30 Denominator = HC01_EST_VC01 Percent = Numerator/denominator denominator_se = HC01_MOE_VC01 numerator_se = $[(HC01\_MOE\_VC30/1.645)^2 + (HC01\_MOE\_VC31/1.645^2)^{0.5}$

			if (numerator_se**2) > ((numerator**2/denominator**2)*((denominator_se)**2)) then pct_se = [sqrt((numerator_se**2) – ((numerator**2/denominator**2)*((denominator_se)**2))/denominator else pct_se = [sqrt((numerator_se**2) + ((numerator**2/denominator**2)*((denominator_se)**2))/denominator
Variable Name	Data Source	Table	Variable(s)
pedshurt	SWITRS	HCI	Note this is the 5-year (2006-2010) annual average rate of severe and fatal pedestrian injuries per 100,000 population Download two files (HCI_RoadTrafficInjuries_753_CT_PL_CO_RE_R4_CA-12-17-13_A-N.zip, HCI_RoadTrafficInjuries_753_CT_PL_CO_12-17-13_O-Y.zip) from https://www.cdph.ca.gov/programs/Pages/HealthyCommunityIndicators.aspx Filter for reportyear (2006-2010), geotype (CT), and mode (pedestrian), copy and paste as .csv then add A-N and O-Y counties into single file (HCI_RoadTrafficInjuriesA-Z.csv) Serious and fatal injuries and rates are added together. R Program does this aggregation se_pct = sqrt(Numerator)/Denominator Note: census tracts with no recorded injuries in SWITRS are set to 0 injuries and 0 injury rate
sealevel	PacInstitute2009		URL: https://archive.cdph.ca.gov/programs/Pages/CalBRACE-Indicators2017.aspx http://pacinst.org/the-impacts-of-sea-level-rise-on-the-california-coast-gis-data-downloads/ Manually filtered BRACE_SLR_784_CT_PL_CO_RE_CA_11-1-2016.xlsx on Total race/ethnicity and census tract geography to create csv file BRACE_Wildfire_CT8-11-2016.csv, which was reformatted in HDI_BRACE_DS_Indicators2017-07-XX.R Percentile calculated only for census tracts with any inundation
traffic	CalEPA	CES 3.0	URL: https://oehha.ca.gov/media/downloads/calenviroscreen/document/ces3results.xlsx traffic = Traffic (Traffic density, in vehicle-kilometers per hour per road length, within 150 meters of the census tract boundary)
transitaccess	HCI2013		UR: https://archive.cdph.ca.gov/programs/Pages/CalBRACE-Indicators2017.aspx Manually filtered RailFerryBus10_MTC_Output_11-15-13.xls, RailFerryBus10_SACOG_Output-11-26-13.xls, RailFerryBus10_SANDAG_Output8-29-13.xls, RailFerryBus10_SCAG_Output9-5-13.xls

			on Total race/ethnicity and census tract geography to create csv file RailFerryBusMTC_SCAG_SACOG_SANDAG2013HCI.csv, which was reformatted in HDI_BRACE_DS_Indicators2017-07-XX.R geotypevalue redefined as text(geotypevalue,"6000000000")
uhii	CalEPA2015		URL: https://calepa.ca.gov/climate/urban-heat-island-index-for-california/ The CalEPA file was segmented by 40 urban location. A SAS file (uhi.sas - N Maizlish 2/2/2016) file was used to combine and de-duplicate the 40 locations into one csv data file. UHII is defined as"positive temperature differential taken every hour over 182 day heat season between an urban census tract and nearby upwind rural reference points at a height of two meters above ground. It units are degree(C^o)-hours. UHII_in_DH_day = UHII/182 Avg_deltaT = UHII_in_DH_day/24 (average temperature difference between an urban census tract and its non-urban reference location
violence	HCI2013		UR: https://archive.cdph.ca.gov/programs/Pages/CalBRACE-Indicators2017.aspx Manually filtered HCI_Crime_752_PL_CO_RE_CA_2000-2013_21OCT15.xlsx for 2013 and place (PL), added FIPS for place code (text(geotypevalue,"60000") to create HCI_CrimePL2013_21OCT15.csv, which was reformatted in R (HDI_BRACE_DS_Indicators2017-07-XX.R)
wildfire	CDFFP2007	FSHZ	URL: https://archive.cdph.ca.gov/programs/Pages/CalBRACE-Indicators2017.aspx http://www.fire.ca.gov/fire_prevention/fire_prevention_wildland_statewide.php Manually filtered BRACE_Wildfire_786_CT_PL_CO_RE_CA.xlsx (downloaded from on Total race/ethnicity and census tract geography to create csv file BRACE_SLR_CT11-1-2016.csv, which was reformatted in HDI_BRACE_DS_Indicators2017-07-XX.R

3. Other Indicators of Disadvantage

Variable Name	Data Source	Table	Variable(s)
ces3score	CalEPA	CES 3.0	URL: https://oehha.ca.gov/media/downloads/calenviroscreen/document/ces3results.xlsx ces3score = CES 3.0 Score
iod	2010 US Census	SF 1 census blocks	URL: http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml (download) Methods follow: https://www.census.gov/prod/2002pubs/censr-3.pdf

hardship	ACS	DP03, DEC_10_DP_DPDP1, DP02, B19301, DP04; S1701	Methods after: Nathan RP, Adams CF. Four Perspectives on Urban Hardship. *Political Science Quarterly.* 1989;104(3):483-508 and Wright DJ, Montiel LM. *Divided They Fall: Hardship in America's Cities and Suburbs*. Albany, NY: The Nelson A. Rockefeller Institute of Government; 2007

C. Data Dictionary for HPI Master Output File (HPI2_MasterFile_13Dec_2017.xlsx)

Variable Name	Definition	Scoring/Comment	Data Source and Year
CensusTract	11-digit census tract code (state+county+tractID)	leading 0 for state is included	
pop2010	total population of census tract in 2010	Decennial Census 2010	
pct2010gq	population in group quarters in 2010	range 0-100	
City	City associated with centroid of census tract	spatially computed field	census_tracts2_place_look-up_table.csv, PHASC
ZIP	5-digit postal zip code	from CES2	CES 3.0
County_FIPS	5-digit code of county	leading 0 for state is not included	2010 Census
County_Name	Name of county	Alameda . . . Yuba	2010 Census
UrbanType	Census classification of urban type	urban, urban_cluster, rural	Decennial Census 2010
hpi2	Healthy Places Index total score (sum of weighted domain scores)	higher number increases advantage	PHASC-VCU, 10/11/17
hpi2_pctile	Healthy Places Index total percentile ranking	0 (most)-100 (least) disadvantaged	PHASC-VCU, 10/11/17
economic	Economic domain score (average of Z-scores of abovepoverty, employed, income), weight 0.319	higher the number, the greater the advantage	PHASC-VCU, 10/11/17
economic_pctile	Economic domain percentile ranking	0 (most)-100 (least) disadvantaged	PHASC-VCU, 10/11/17
education	Education domain score (average of Z-scores of bachelorsed,inpreschool, inhischool), weight 0.187	higher number increases advantage	PHASC-VCU, 10/11/17
education_pctile	Education domain percentile ranking	0 (most)-100 (least) disadvantaged	PHASC-VCU, 10/11/17
housing	Housing domain score (average of Z-scores of uncrowded, ownsever,rentsever,homeownership), weight 0.052	higher number increases advantage	PHASC-VCU, 10/11/17
housing_pctile	Housing domain percentile ranking	0 (most)-100 (least) disadvantaged	PHASC-VCU, 10/11/17
insured	Healthcare Access domain score (Z-score of insurance), weight 0.052	higher number increases advantage	PHASC-VCU, 10/11/17
insured_pctile	Healthcare Access domain percentile ranking	0 (most)-100 (least) disadvantaged	PHASC-VCU, 10/11/17
neighborhood	Neighborhood domain score (average of Z-scores of parksaccess, retail, treecanopy,supermrkts, alcoffsale), weight 0.077	higher number increases advantage	PHASC-VCU, 10/11/17

neighborhood_pctile	Neighborhood domain percentile ranking	0 (most)-100 (least) disadvantaged	PHASC-VCU, 10/11/17
pollution	Clean Environment domain score (average of Z-scores of dieslpm, ozone, pm25, h20contam), weight 0.052	higher number increases advantage	PHASC-VCU, 10/11/17
pollution_pctile	Clean Environment domain percentile ranking	0 (most)-100 (least) disadvantaged	PHASC-VCU, 10/11/17
transportation	Transportation domain score (average of Z-scores of commute, automobile), weight 0.155	higher number increases advantage	PHASC-VCU, 10/11/17
transportation_pctile	Transportation domain percentile ranking	0 (most)-100 (least) disadvantaged	PHASC-VCU, 10/11/17
social	Social domain score (average of Z-scores of twoparents, voting), weight 0.104	higher number increases advantage	PHASC-VCU, 10/11/17
social_pctile	Social domain percentile ranking	0 (most)-100 (least) disadvantaged	PHASC-VCU, 10/11/17
abovepoverty	Percent of the population with an income exceeding 200% of federal poverty level (value - %)	higher number increases advantage	ACS, 2011-2015
abovepoverty_pctile	Percent of the population with an income exceeding 200% of federal poverty level (percentile ranking)	0 (most)-100 (least) disadvantaged	
alcoffsale	Percentage of the population residing beyond ¼ mile of an off-site sales alcohol outlet (value - %)	higher number increases advantage	ABC, 2011-2015
alcoffsale_pctile	Percentage of the population residing beyond ¼ mile of an off-site sales alcohol outlet (percentile ranking)	0 (most)-100 (least) disadvantaged	
automobile	Percentage of households with access to an automobile (value - %)	higher number increases advantage	ACS, 2011-2015
automobile_pctile	Percentage of households with access to an automobile (percentile ranking)	0 (most)-100 (least) disadvantaged	
bachelorsed	Percentage of population over age 25 with a bachelor's education or higher (value - %)	higher number increases advantage	ACS, 2011-2015
bachelorsed_pctile	Percentage of population over age 25 with a bachelor's education or higher (percentile ranking)	0 (most)-100 (least) disadvantaged	
commute	Percentage of workers (16 years and older) who commute to work by transit, walking, or cycling (value - %)	higher number increases advantage	ACS, 2011-2015
commute_pctile	Percentage of workers (16 years and older) who commute to work by transit, walking, or cycling (percentile ranking)	0 (most)-100 (least) disadvantaged	
dieselpm	Spatial distribution of gridded diesel PM emissions from on-road and non-road sources for a 2012 summer day in July (kg/day) (value - %)	**higher number decreases advantage**	CalEPA, 2012
dieselpm_pctile	Spatial distribution of gridded diesel PM emissions from on-road and non-road sources for a 2012 summer day in July (kg/day) (percentile ranking)	0 (most)-100 (least) disadvantaged	

employed	Percentage of population aged 25-64 who are employed (value - %)	higher number increases advantage	ACS, 2011-2015
employed_pctile	Percentage of population aged 25-64 who are employed (percentile ranking)	0 (most)-100 (least) disadvantaged	
h20contam	CalEnviroScreen 3.0 drinking water contaminant index for selected contaminants (value - %)	**higher number decreases advantage**	CalEPA, 2011-2013
h20contam_pctile	CalEnviroScreen 3.0 drinking water contaminant index for selected contaminants (percentile ranking)	0 (most)-100 (least) disadvantaged	
homeownership	Percentage of occupied housing units occupied by property owners (value - %)	higher number increases advantage	ACS, 2011-2015
homeownership_pctile	Percentage of occupied housing units occupied by property owners (percentile ranking)	0 (most)-100 (least) disadvantaged	
houserepair	Percent of households with kitchen facilities and plumbing (value - %)	higher number increases advantage	ACS, 2011-2015
houserepair_pctile	Percent of households with kitchen facilities and plumbing (percentile ranking)	0 (most)-100 (least) disadvantaged	
income	Median annual household income (value - %)	higher number increases advantage	ACS, 2011-2015
income_pctile	Median annual household income (percentile ranking)	0 (most)-100 (least) disadvantaged	
inhighschool	Percentage of 15-17 year olds enrolled in school (value - %)	higher number increases advantage	ACS, 2011-2015
inhighschool_pctile	Percentage of 15-17 year olds enrolled in school (percentile ranking)	0 (most)-100 (least) disadvantaged	
inpreschool	Percentage of 3 and 4 year olds enrolled in school (value - %)	higher number increases advantage	ACS, 2011-2015
inpreschool_pctile	Percentage of 3 and 4 year olds enrolled in school (percentile ranking)	0 (most)-100 (least) disadvantaged	
insurance	Percentage of adults aged 18 to 64 years currently insured (value - %)	higher number increases advantage	ACS, 2011-2015
insurance_pctile	Percentage of adults aged 18 to 64 years currently insured (percentile ranking)	0 (most)-100 (least) disadvantaged	
ownsevere	Percentage of low-income homeowners paying more than 50% of income on housing costs (value - %)	**higher number decreases advantage**	CHAS, 2010-2014
ownsevere_pctile	Percentage of low-income homeowners paying more than 50% of income on housing costs (percentile ranking)	0 (most)-100 (least) disadvantaged	
ozone	Mean of summer months (May-October) of the daily maximum 8-hour ozone concentration (ppm) averaged over three years (2012 to 2014) (value - %)	**higher number decreases advantage**	CalEPA, 2012-2014
ozone_pctile	Mean of summer months (May-October) of the daily maximum 8-hour ozone concentration (ppm) averaged over three years (2012 to 2014) (percentile ranking)	0 (most)-100 (least) disadvantaged	

parkaccess	Percentage of the population living within a half-mile of a park, beach, or open space greater than 1 acre (value - %)	higher number increases advantage	GreenInfo, 2012
parkaccess_pctile	Percentage of the population living within a half-mile of a park, beach, or open space greater than 1 acre (percentile ranking)	0 (most)-100 (least) disadvantaged	
pm25	Annual mean concentration of PM2.5 (average of quarterly means, µg/m3), over three years (2012 to 2014) (value - %)	**higher number decreases advantage**	CalEPA, 2012-2014
pm25_pctile	Annual mean concentration of PM2.5 (average of quarterly means, µg/m3), over three years (2012 to 2014) (percentile ranking)	0 (most)-100 (least) disadvantaged	
rentsevere	Percentage of low income renter households paying more than 50% of income on housing costs (value - %)	**higher number decreases advantage**	CHAS, 2010-2014
rentsevere_pctile	Percentage of low income renter households paying more than 50% of income on housing costs (percentile ranking)	0 (most)-100 (least) disadvantaged	
retail	Gross retail, entertainment, and education employment density (jobs/acre) on unprotected land (value - %)	higher number increases advantage	USEPA, 2010
retail_pctile	Gross retail, entertainment, and education employment density (jobs/acre) on unprotected land (percentile ranking)	0 (most)-100 (least) disadvantaged	
supermkts	Percentage of the urban population residing less than 1/2 mile from a supermarket/large grocery store, or the percent of the rural population living less than 1 miles from a supermarket/large grocery store (value - %)	higher number increases advantage	USDA, 2015
supermkts_pctile	Percentage of the urban population residing less than 1/2 mile from a supermarket/large grocery store, or the percent of the rural population living less than 1 miles from a supermarket/large grocery store (percentile ranking)	0 (most)-100 (least) disadvantaged	
treecanopy	Population-weighted percentage of the census tract area with tree canopy (value - %)	higher number increases advantage	NLCD, 2011
treecanopy_pctile	Population-weighted percentage of the census tract area with tree canopy (percentile ranking)	0 (most)-100 (least) disadvantaged	
twoparents	Percentage of children in married-couple family households or one parent with unmarried partner family households (value - %)	higher number increases advantage	ACS, 2011-2015
twoparents_pctile	Percentage of children in married-couple family households or one parent with unmarried partner family households (percentile ranking)	0 (most)-100 (least) disadvantaged	

uncrowded	Percentage of households with less or equal to 1 occupant per room (value - %)	higher number increases advantage	ACS, 2011-2015
uncrowded_pctile	Percentage of households with less or equal to 1 occupant per room (percentile ranking)	0 (most)-100 (least) disadvantaged	
voting	Percentage of registered voters who voted in the 2012 general election (value - %)	higher number increases advantage	UC Berkeley, 2012
voting_pctile	Percentage of registered voters who voted in the 2012 general election (percentile ranking)	0 (most)-100 (least) disadvantaged	
LEB	Life Expectancy at Birth, 2010	64.8-90.0 (least-most advantaged)	VCU2010
LEB_pctile	Percentile ranking of LEB	0-100 (least-most advantaged)	
asian_pct	Percent of Asians in the total population	0-100	DEC_10_SF2_PCT1
black_pct	Percent of Blacks in the total population	0-100	DEC_10_SF2_PCT1
latino_pct	Percent of Latinos in the total population	0-100	DEC_10_SF2_PCT1
multiple_pct	Percent of two or more races in the total population	0-100	DEC_10_SF2_PCT1
NativeAm_pct	Percent of American Indian/Alaskan Natives in the total population	0-100	DEC_10_SF2_PCT1
other_pct	Percent of some other race in the total population	0-100	DEC_10_SF2_PCT1
PacificIsl_pct	Percent of Native Hawaiians and other Pacific Islanders in the total population	0-100	DEC_10_SF2_PCT1
white_pct	Percent of Whites in the total population	0-100	DEC_10_SF2_PCT1
version	Date file was created	Day of Week, Month Day Year	PHASC/VCU

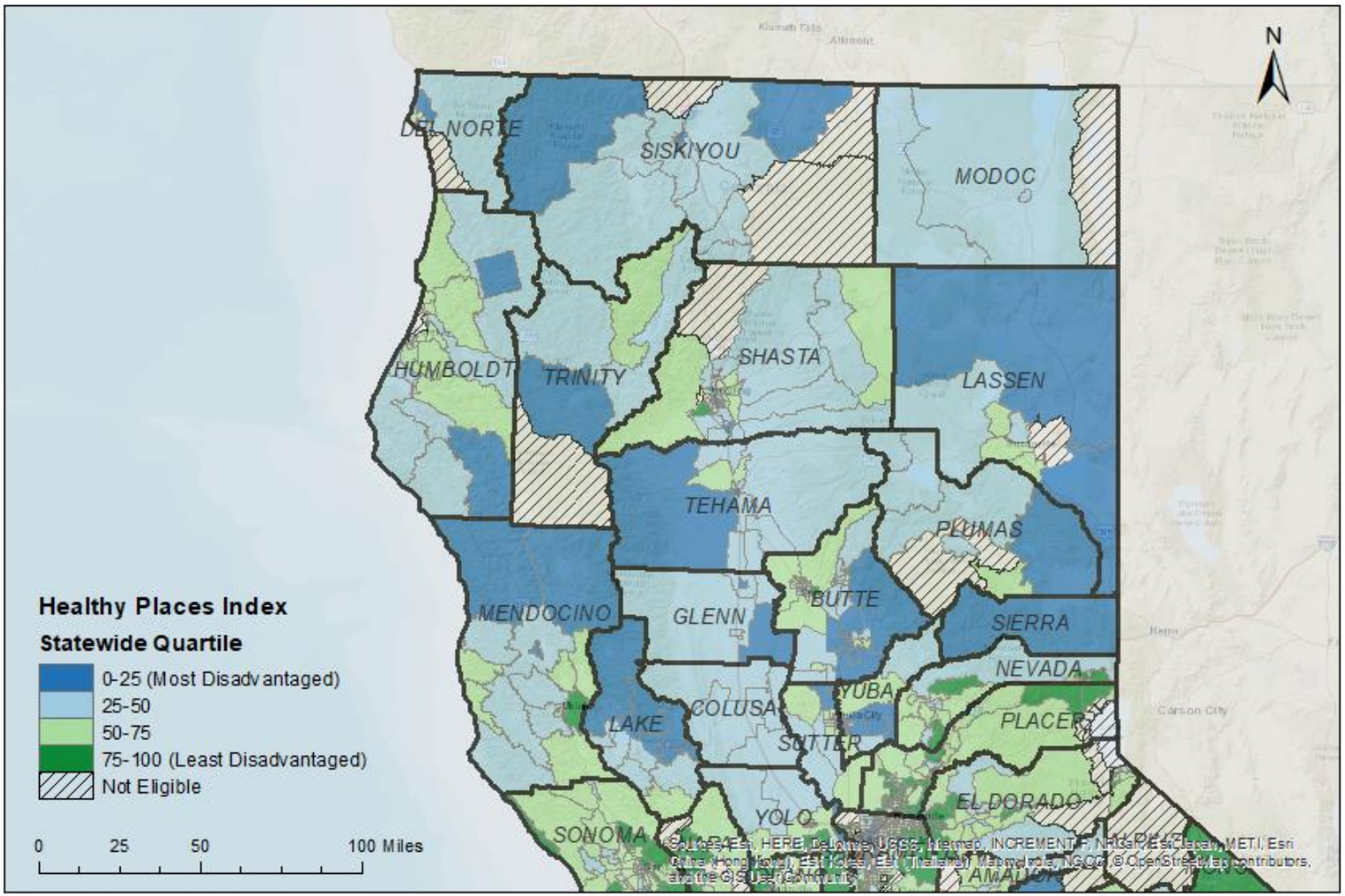
ABC, Alcoholic Beverage Commission; ACS, American Community Survey; CHAS, Comprehensive Housing Assessment System; CalEPA; California Environmental Protection Agency; NLCD, National Land Cover Database; USDA FARA, U.S. Department of Agriculture Food Access Research Atlas; USEPA, U.S. Environmental Protection Agency; UC Berkeley, University of California, Berkeley; VCU, Virginia Commonwealth University

D. Maps of the Healthy Places Index by California Census Tract

public health alliance of southern california

DRAFT - 11/3/2017

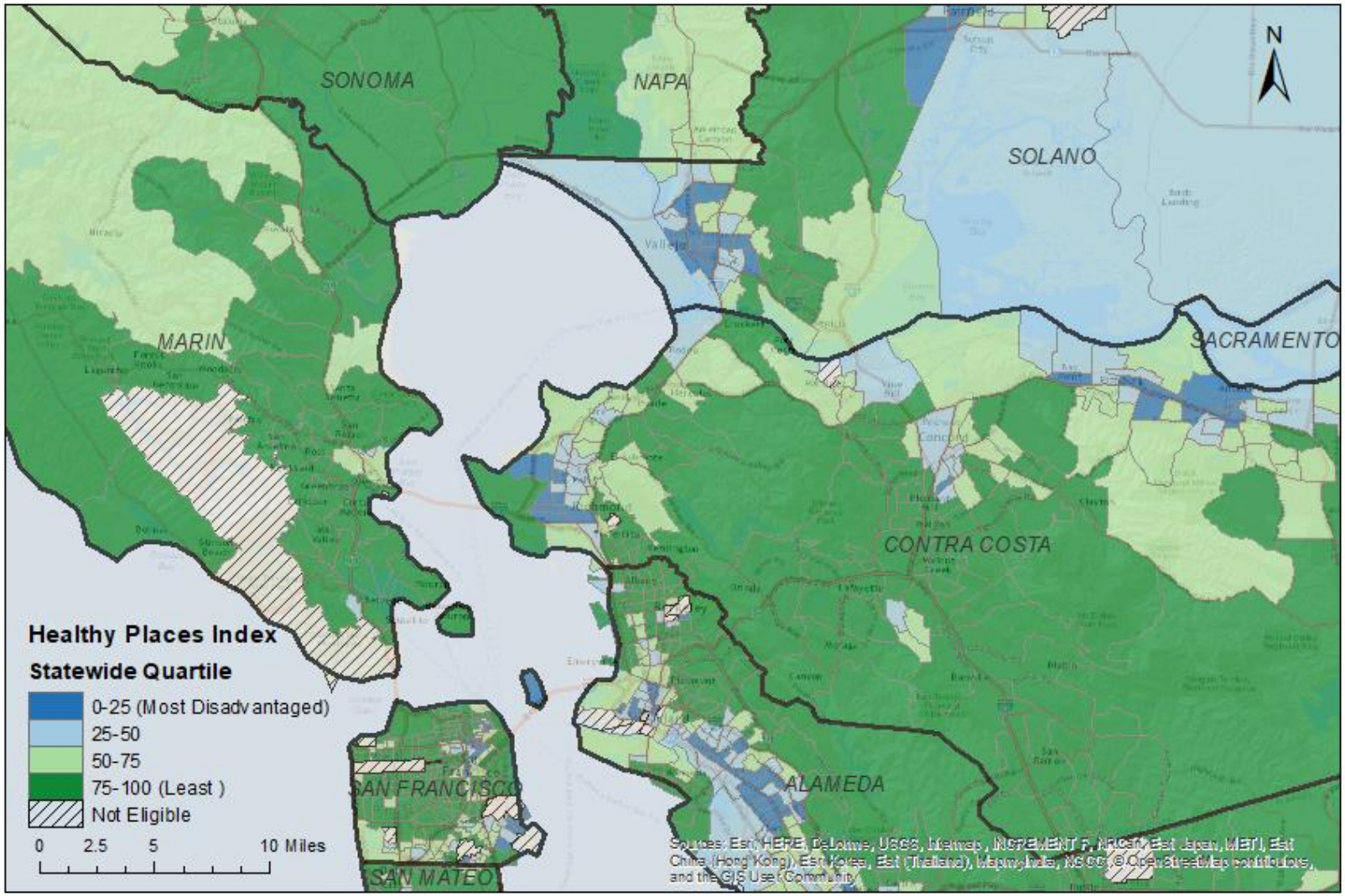
Census Tract Distribution of Healthy Places Index by Quartiles, Northern California



public health alliance of southern california

DRAFT - 11/3/2017

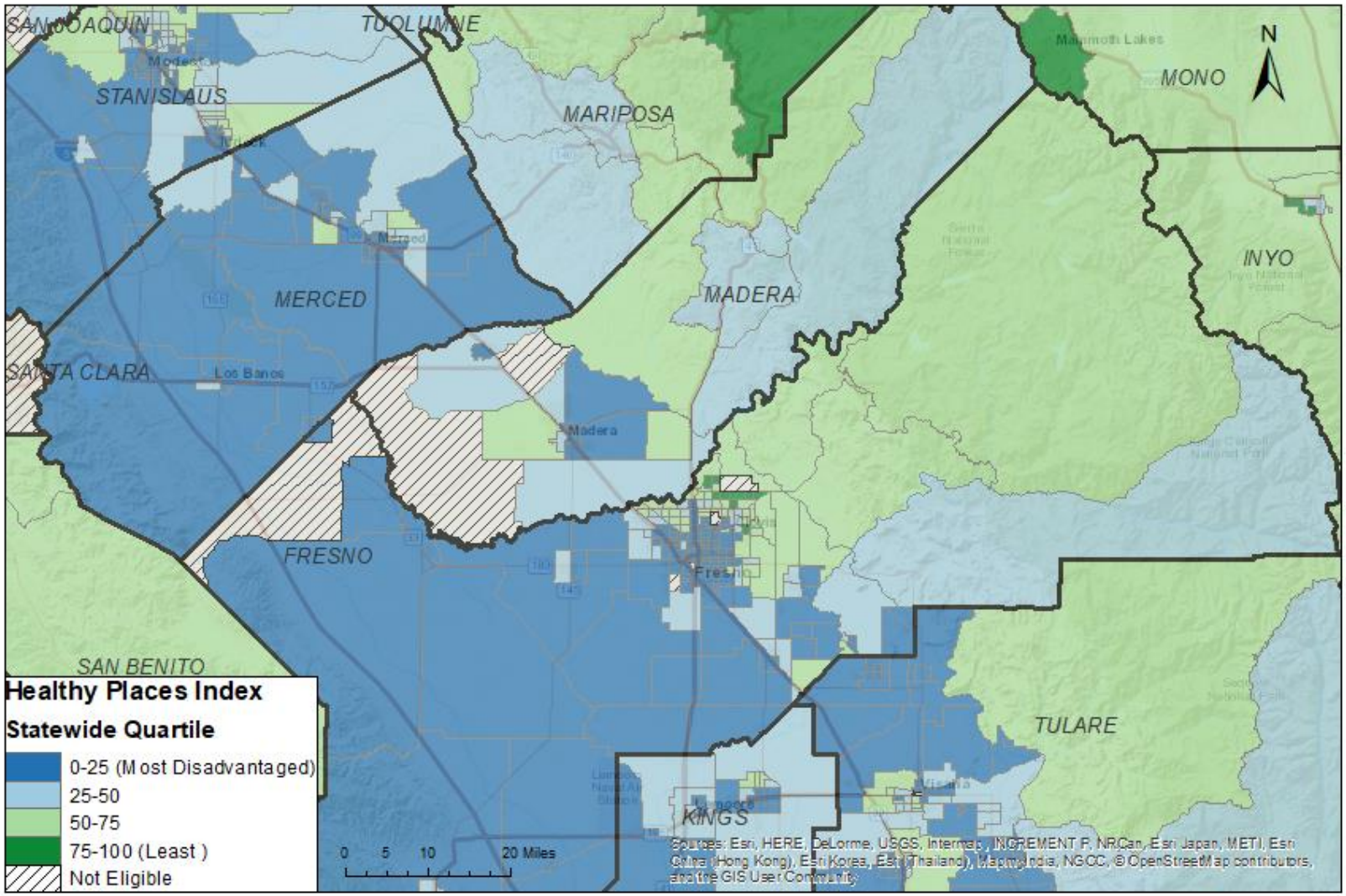
Census Tract Distribution of Healthy Places Index by Quartiles, SF Bay Area



public health alliance of southern california

DRAFT - 11/3/2017

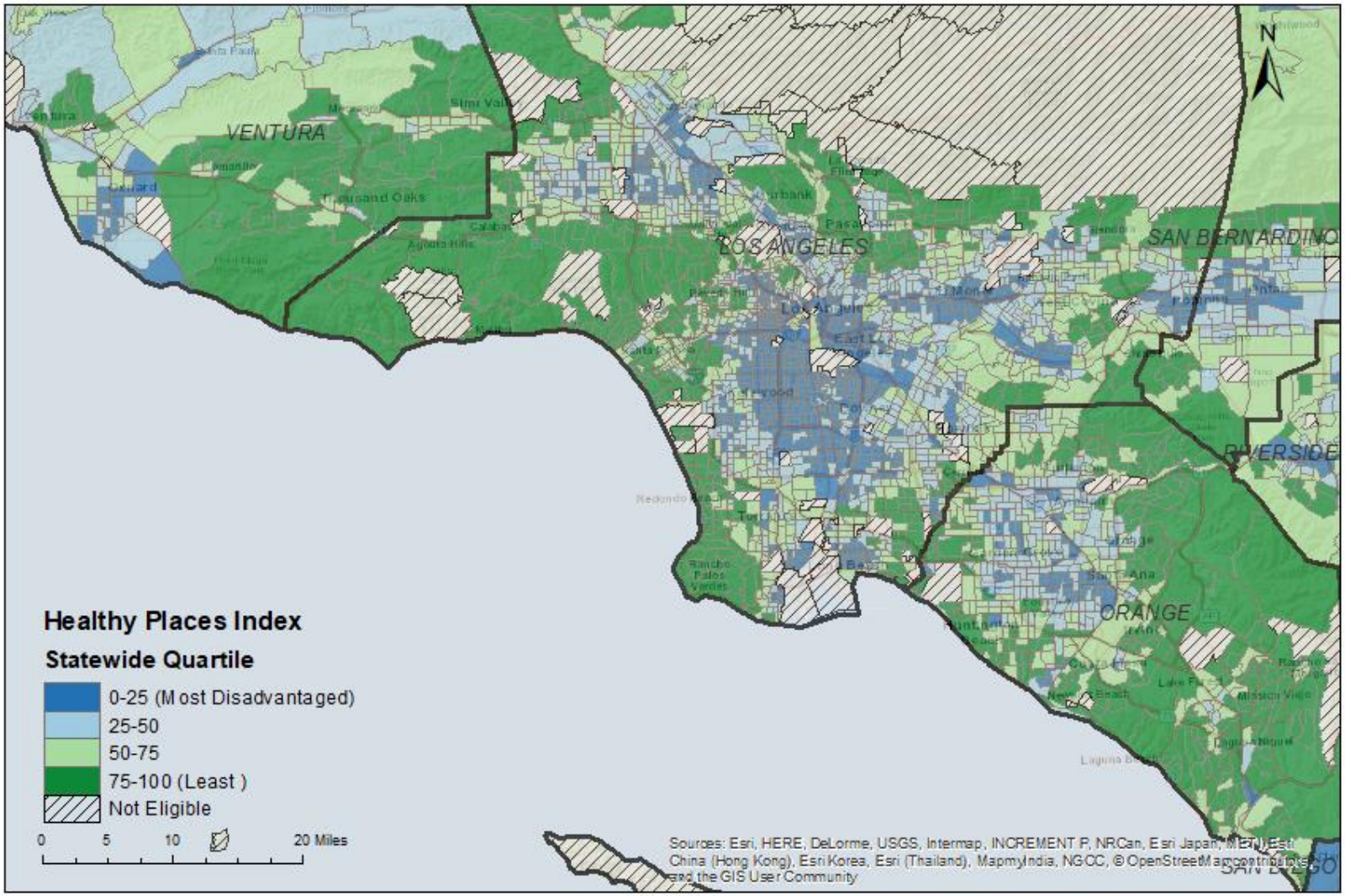
Census Tract Distribution of Healthy Places Index by Quartiles, Mid-Central Valley



public health alliance of southern california

DRAFT - 11/3/2017

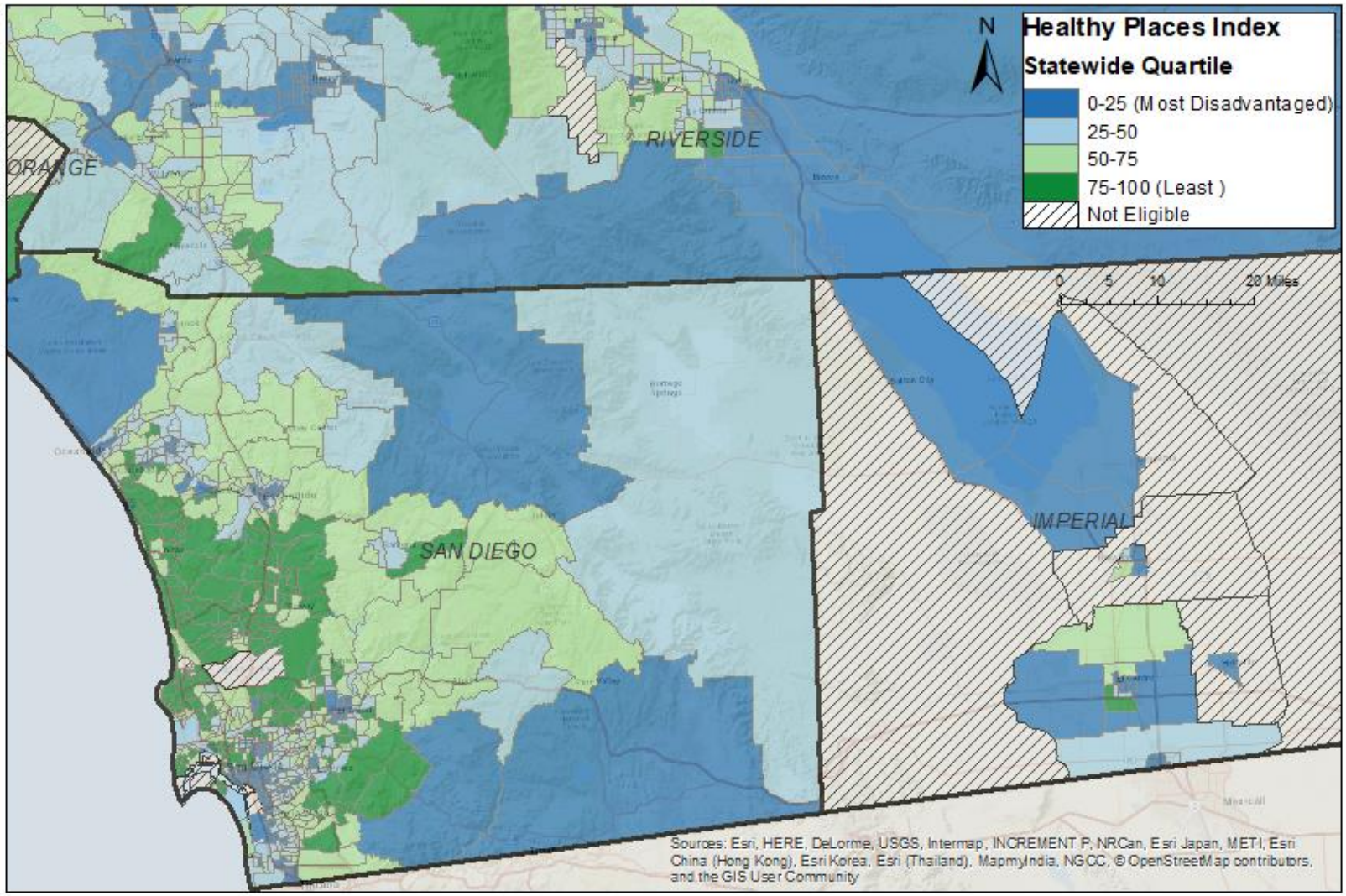
Census Tract Distribution of Healthy Places Index by Quartiles, Los Angeles Area



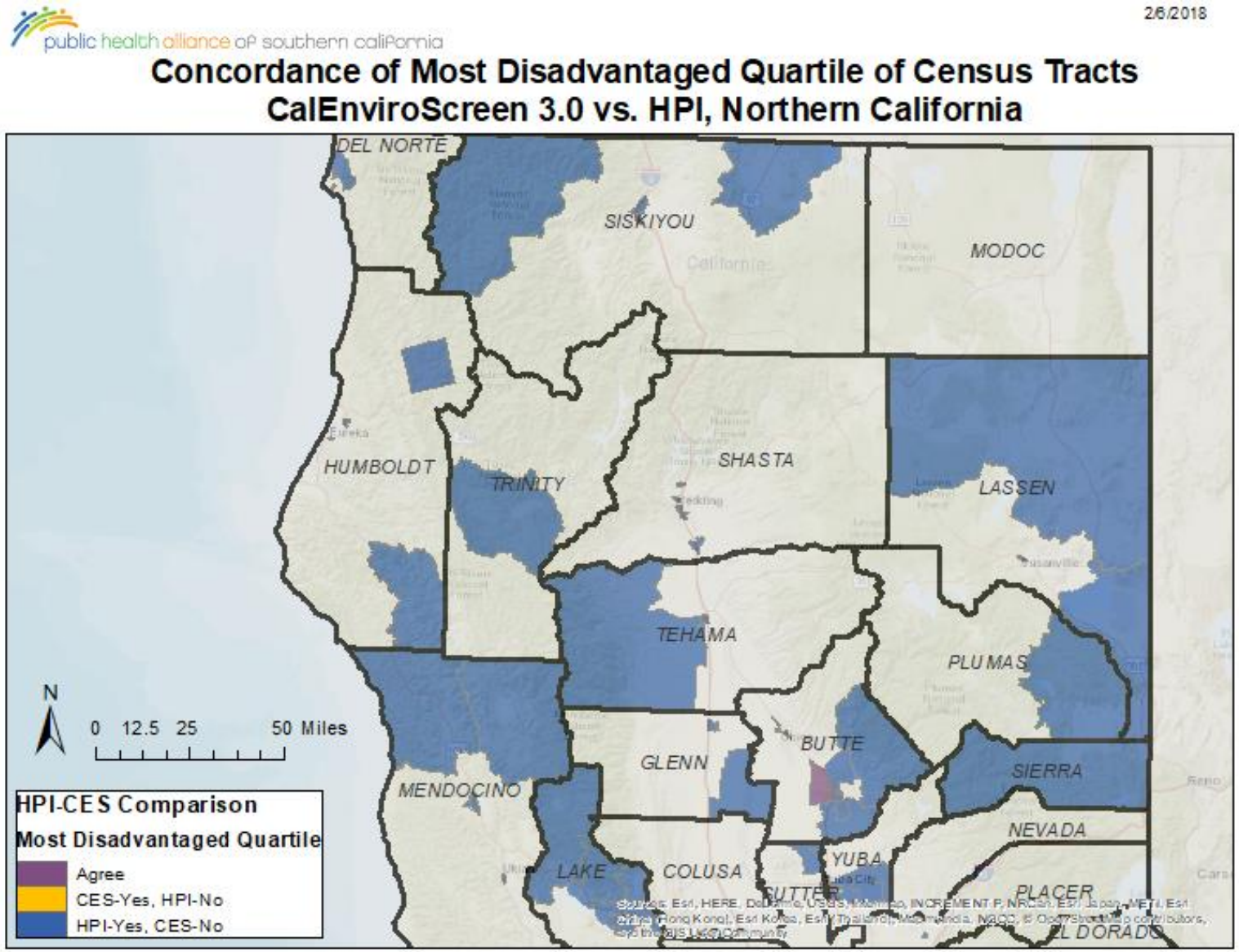
public health alliance of southern california

DRAFT - 11/3/2017

Census Tract Distribution of Healthy Places Index by Quartiles, San Diego, Imperial, and Riverside Counties



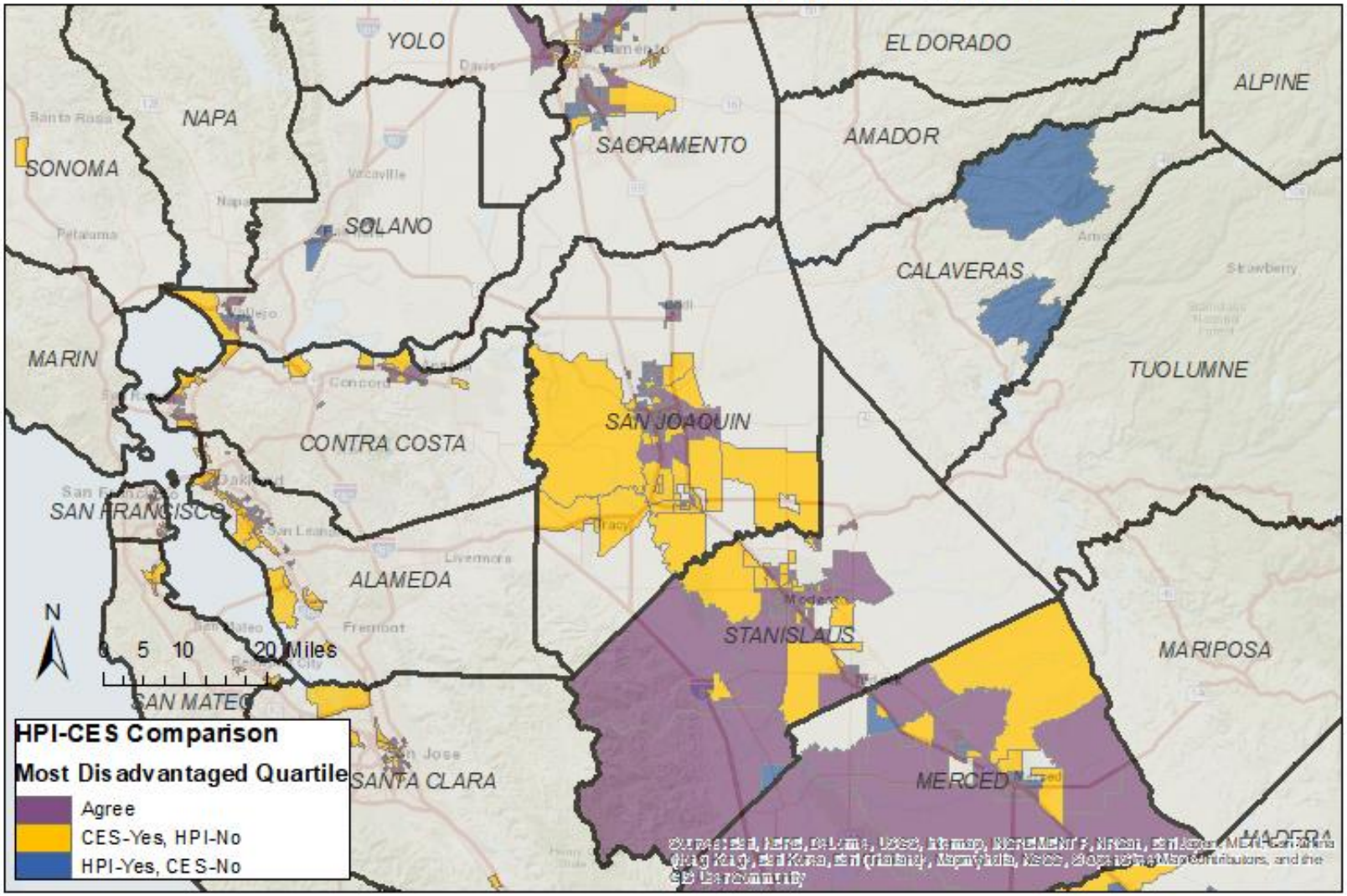
Concordance of Most Disadvantaged Quartile of Census Tracts: CalEnviroScreen 3.0 vs. Healthy Places Index



public health alliance of southern california

2/6/2018

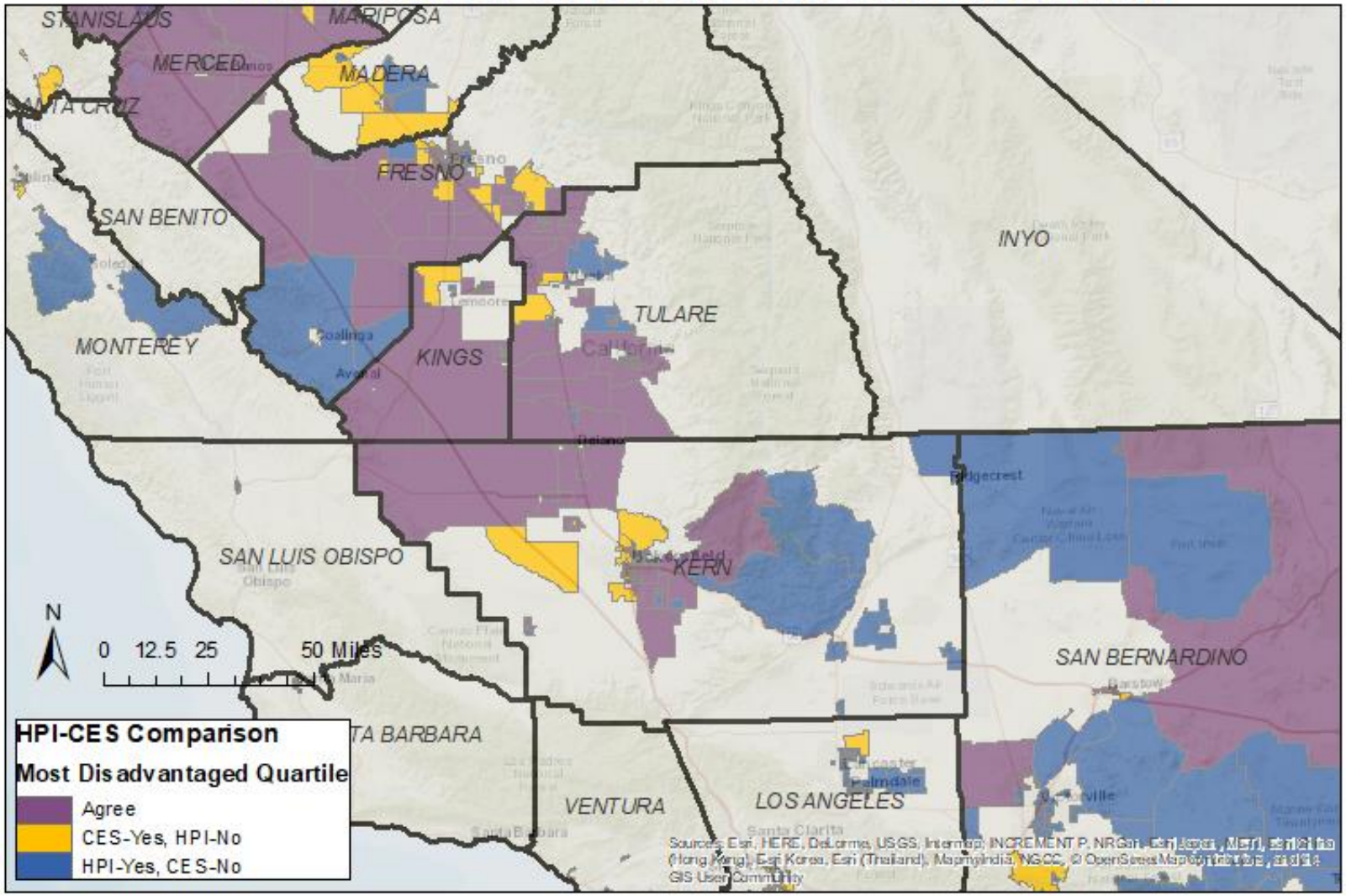
Concordance of Most Disadvantaged Quartile of Census Tracts
CalEnviroScreen 3.0 vs. HPI, Mid-California



public health alliance of southern california

2/6/2018

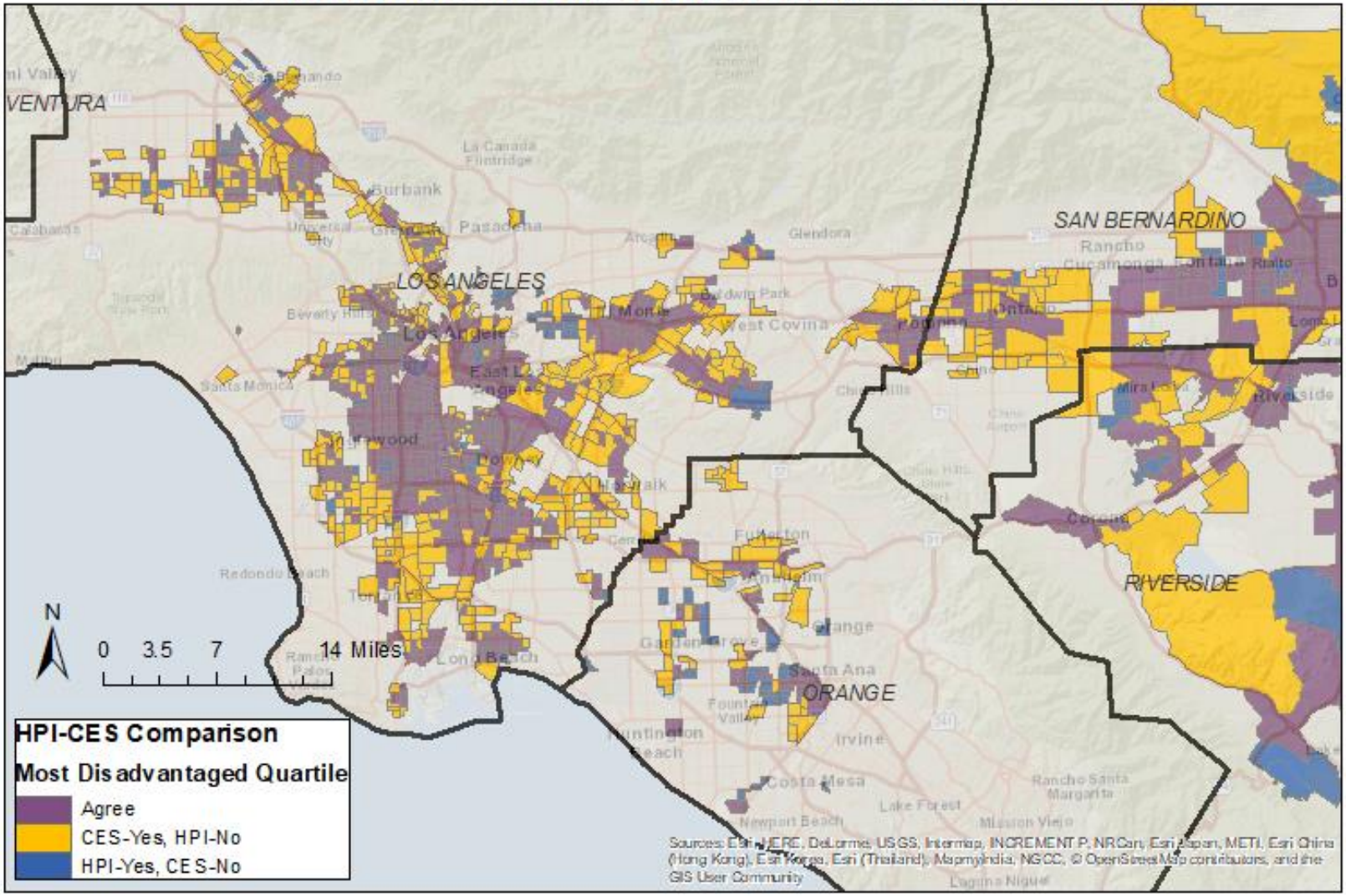
Concordance of Most Disadvantaged Quartile of Census Tracts
CalEnviroScreen 3.0 vs. HPI, Lower San Joaquin Valley



public health alliance of southern california

2/6/2018

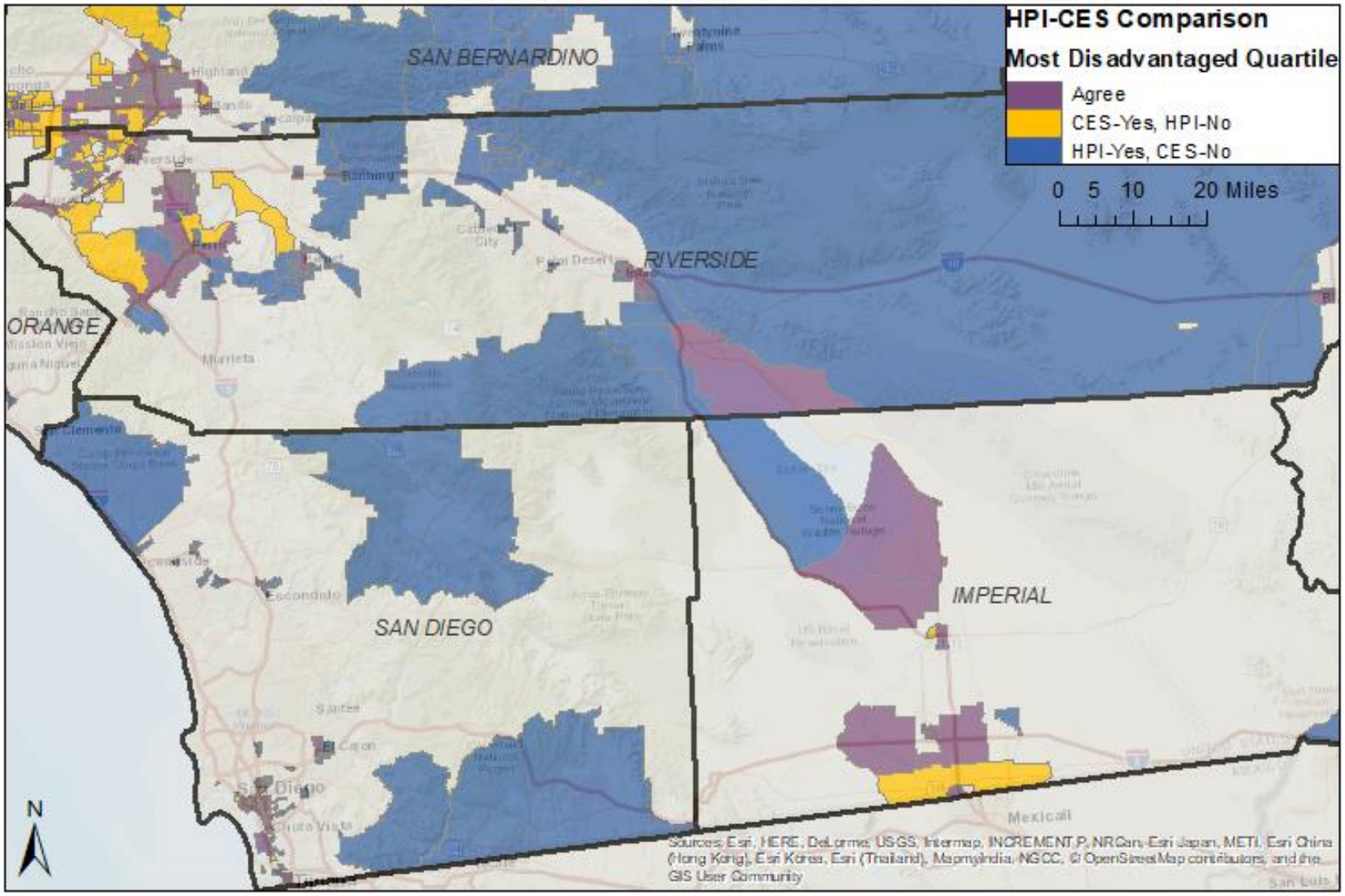
Concordance of Most Disadvantaged Quartile of Census Tracts
CalEnviroScreen 3.0 vs. HPI, Los Angeles Area Counties



public health alliance of southern california

2/6/2018

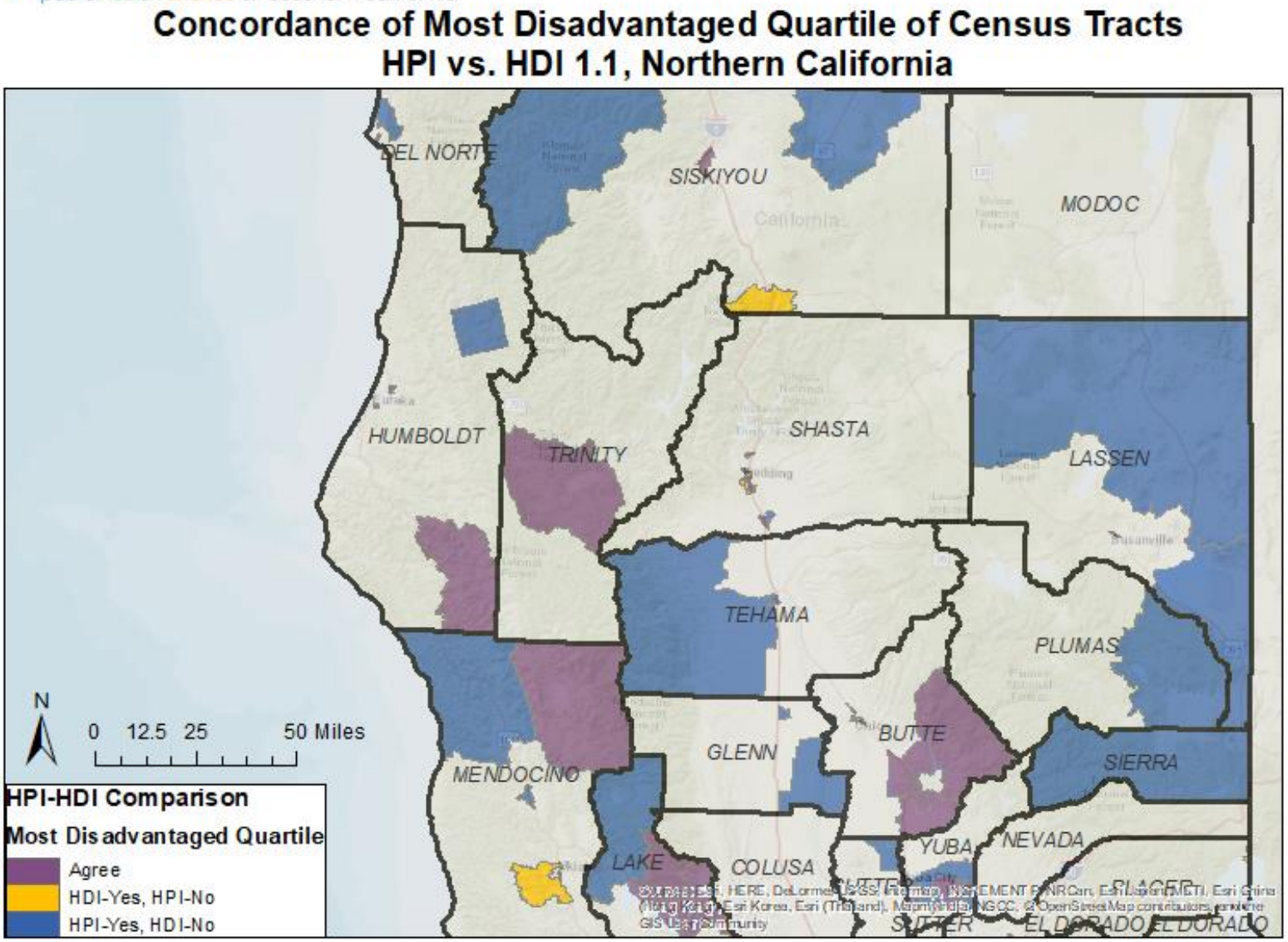
Concordance of Most Disadvantaged Quartile of Census Tracts CalEnviroScreen 3.0 vs. HPI, Southern California



Concordance of Most Disadvantaged Quartile of Census Tracts: Healthy Places Index vs. Health Disadvantage Index 1.1

public health alliance of southern california

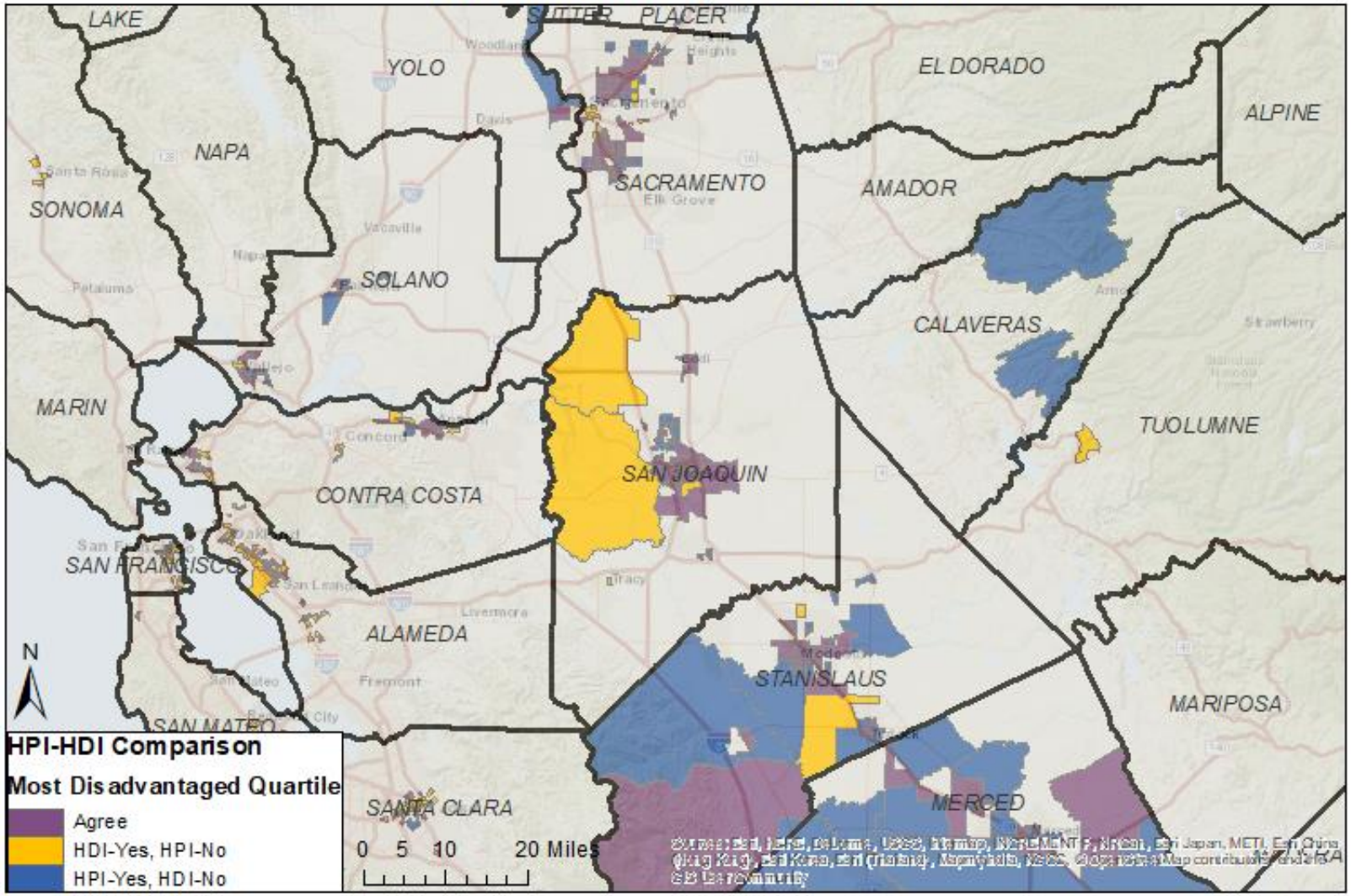
2/6/2018



public health alliance of southern california

2/6/2018

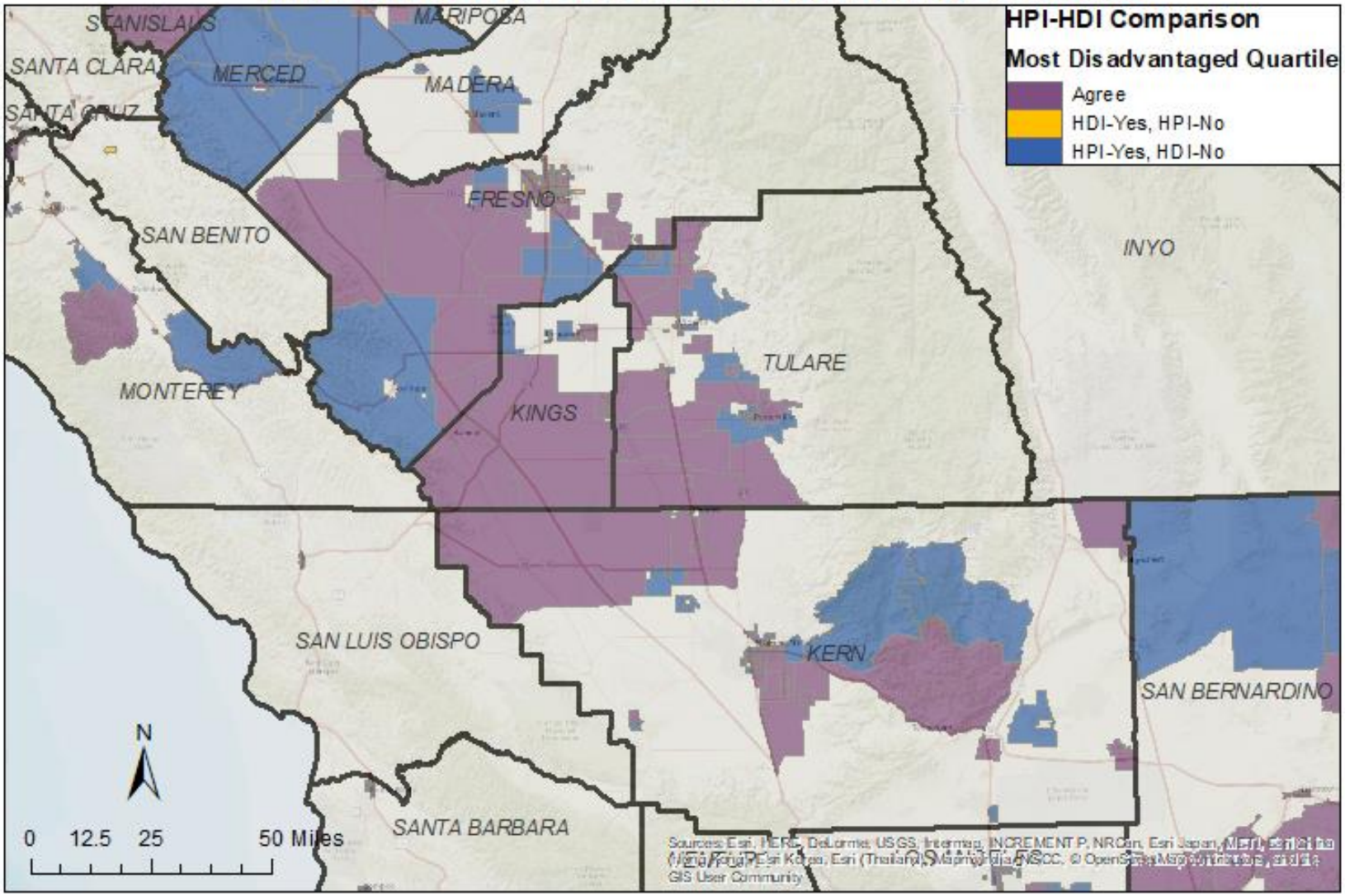
Concordance of Most Disadvantaged Quartile of Census Tracts
HPI vs. HDI 1.1, Mid-California



public health alliance of southern california

2/6/2018

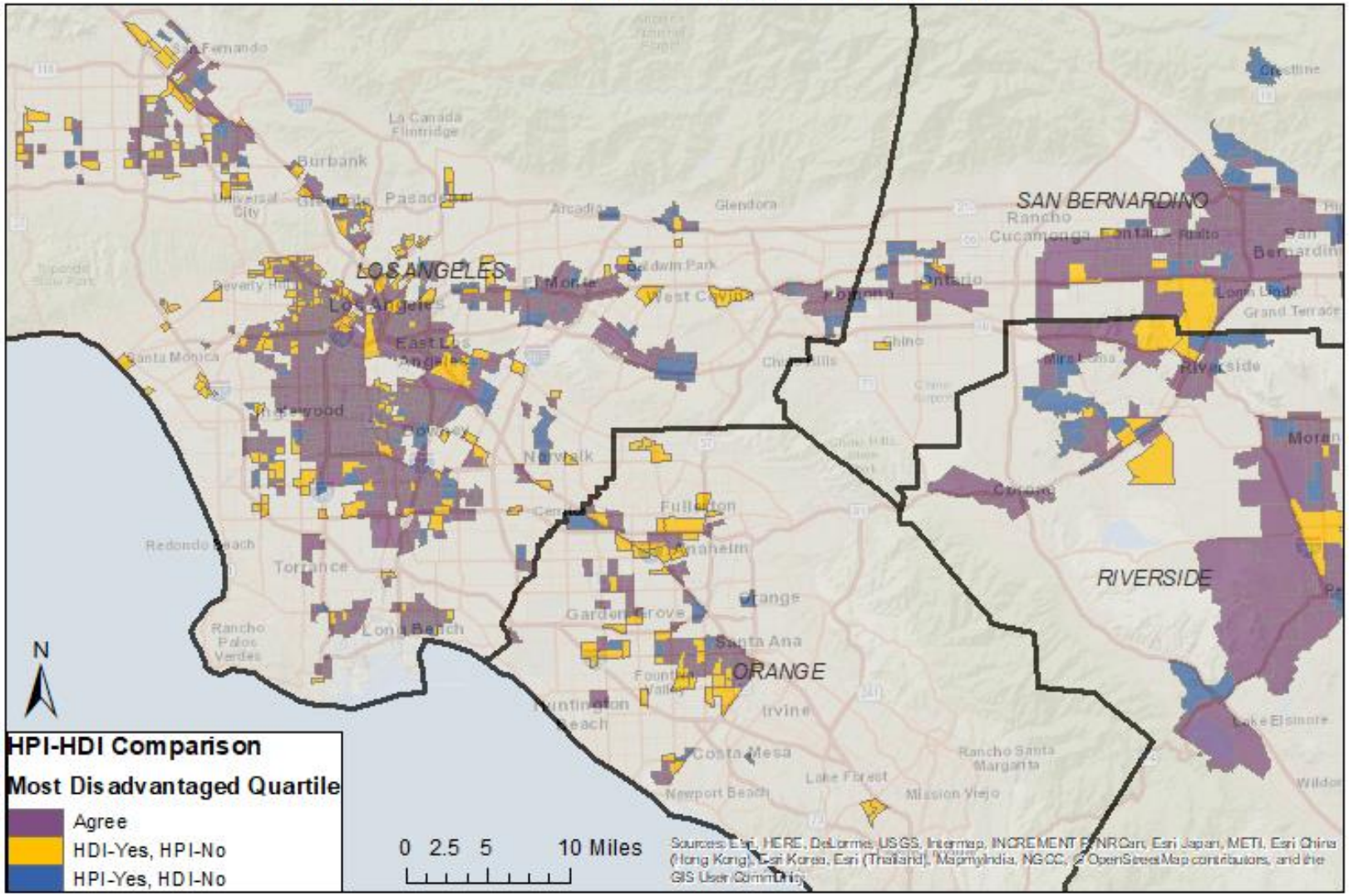
Concordance of Most Disadvantaged Quartile of Census Tracts
HPI vs. HDI 1.1, Lower San Joaquin Valley Counties



public health alliance of southern california

2/6/2018

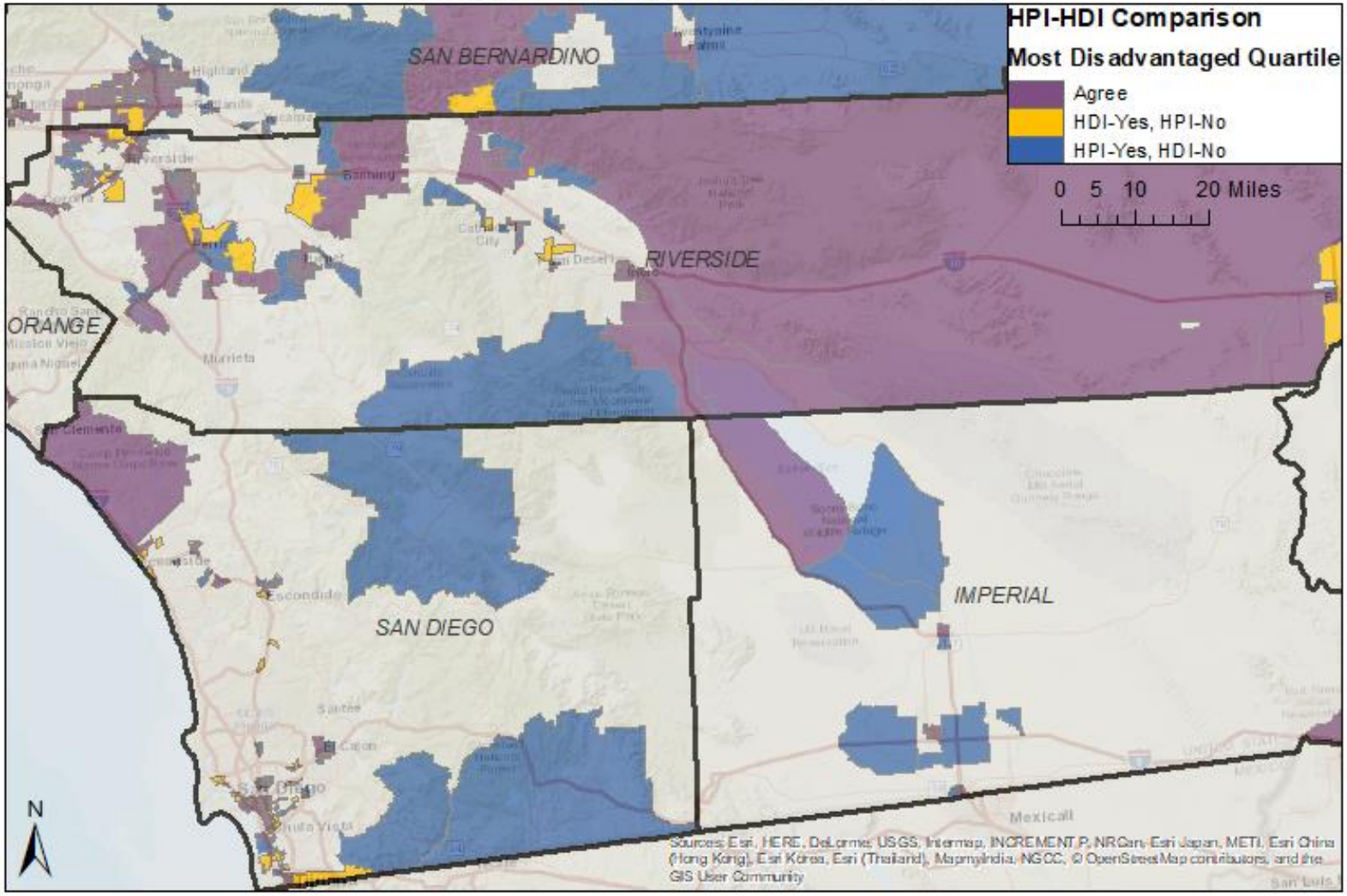
Concordance of Most Disadvantaged Quartile of Census Tracts
HPI vs. HDI 1.1, Los Angeles Area Counties



public health alliance of southern california

2/6/2018

Concordance of Most Disadvantaged Quartile of Census Tracts HDI1.1 vs. HPI, Southern California

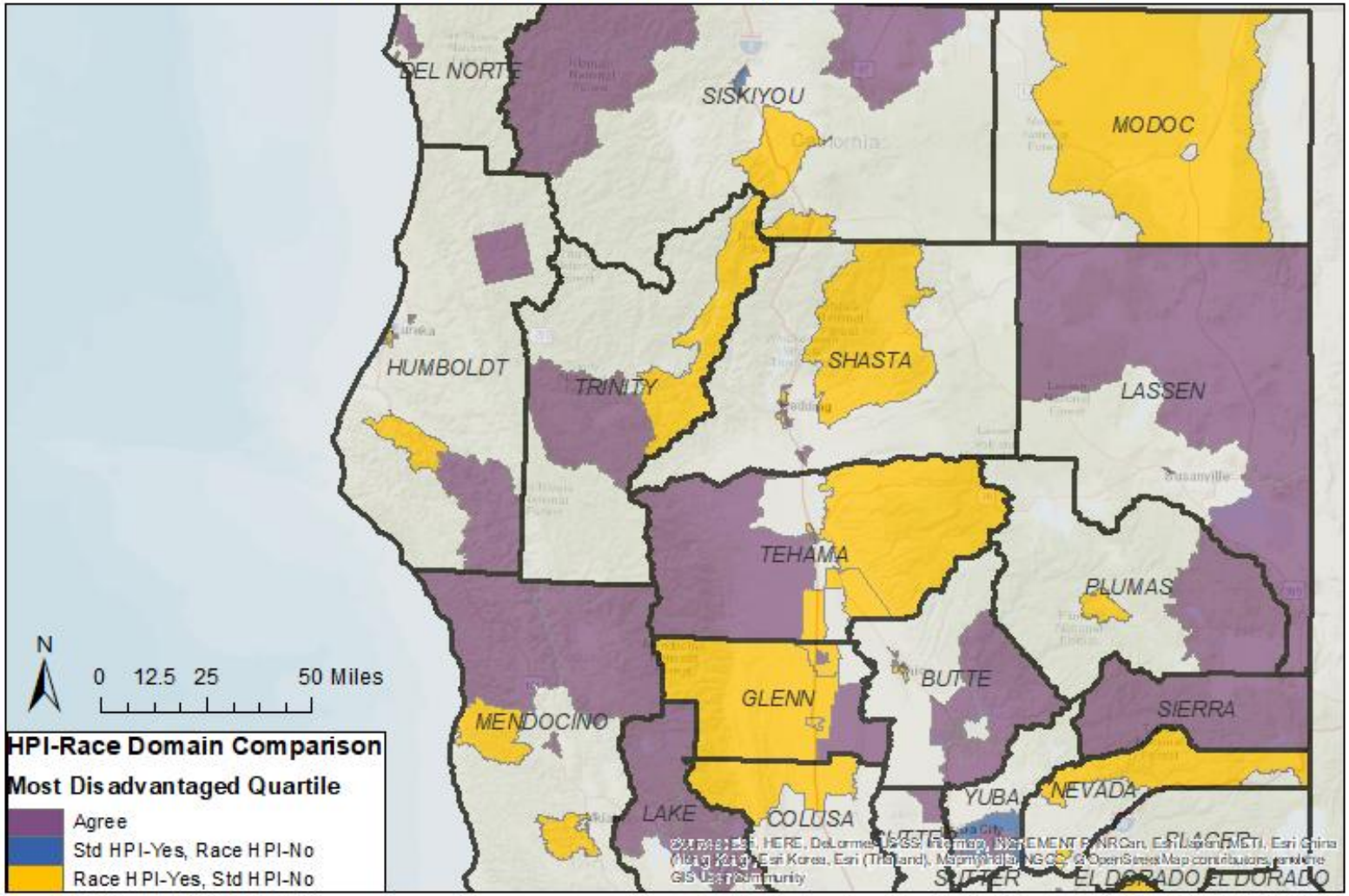


Concordance of Most Disadvantaged Quartile of Census Tracts: Healthy Places Index With and Without Race Domain

2/6/2018

public health alliance of southern california

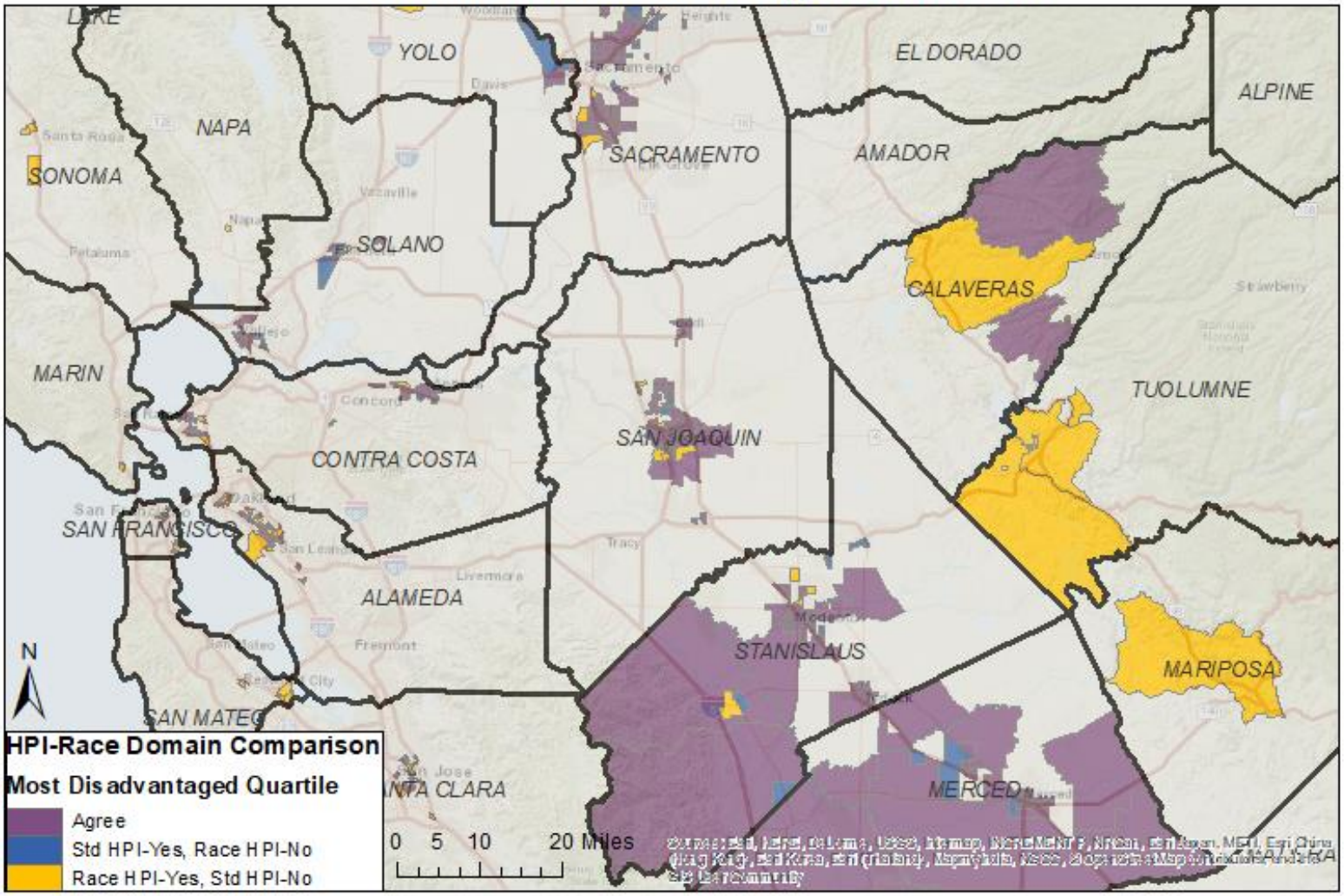
Concordance of Most Disadvantaged Quartile of Census Tracts
Standard HPI vs. Race HPI, Northern California



public health alliance of southern california

2/6/2018

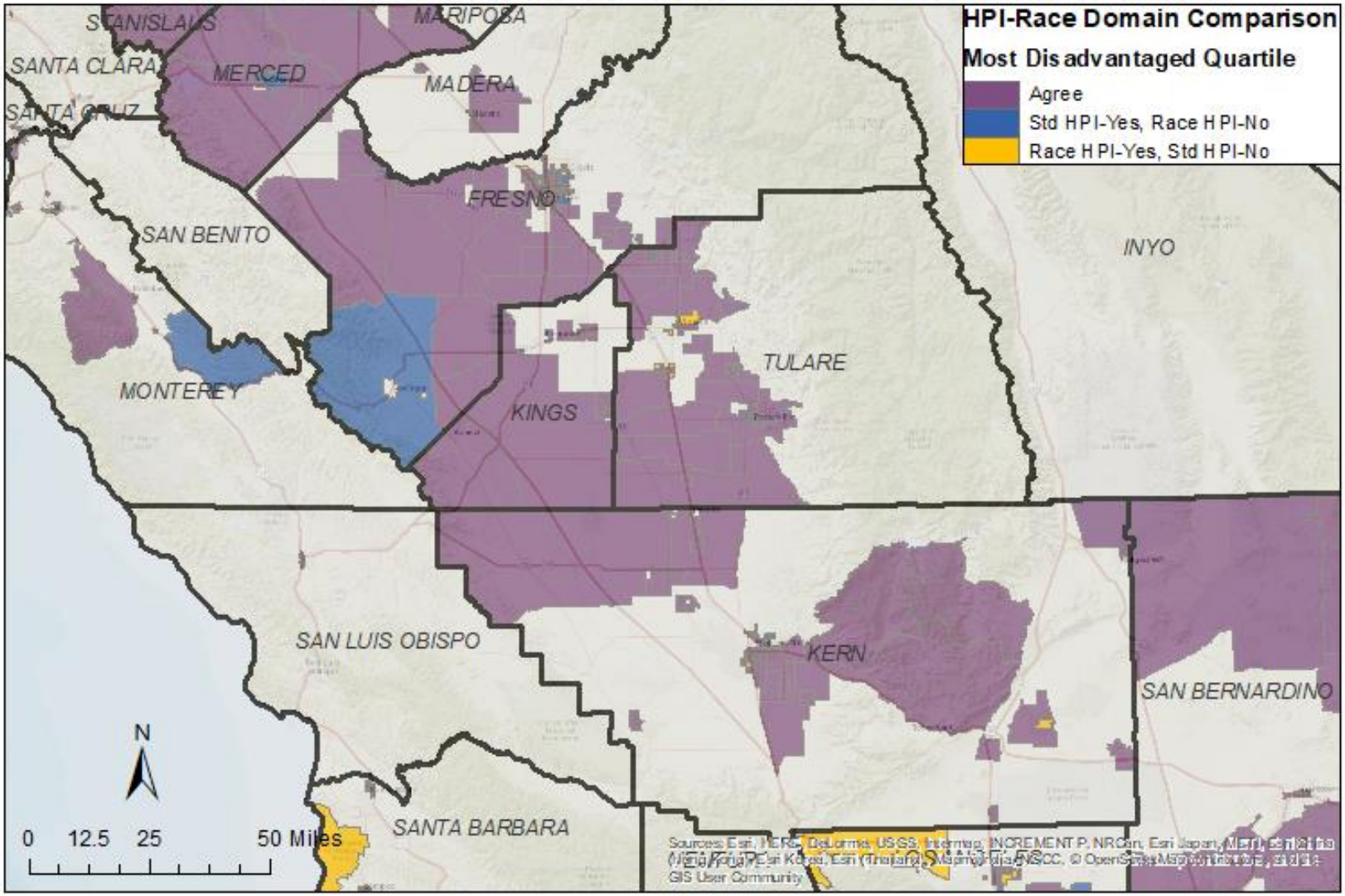
Concordance of Most Disadvantaged Quartile of Census Tracts
Standard HPI vs. Race HPI, Mid-California



public health alliance of southern california

2/6/2018

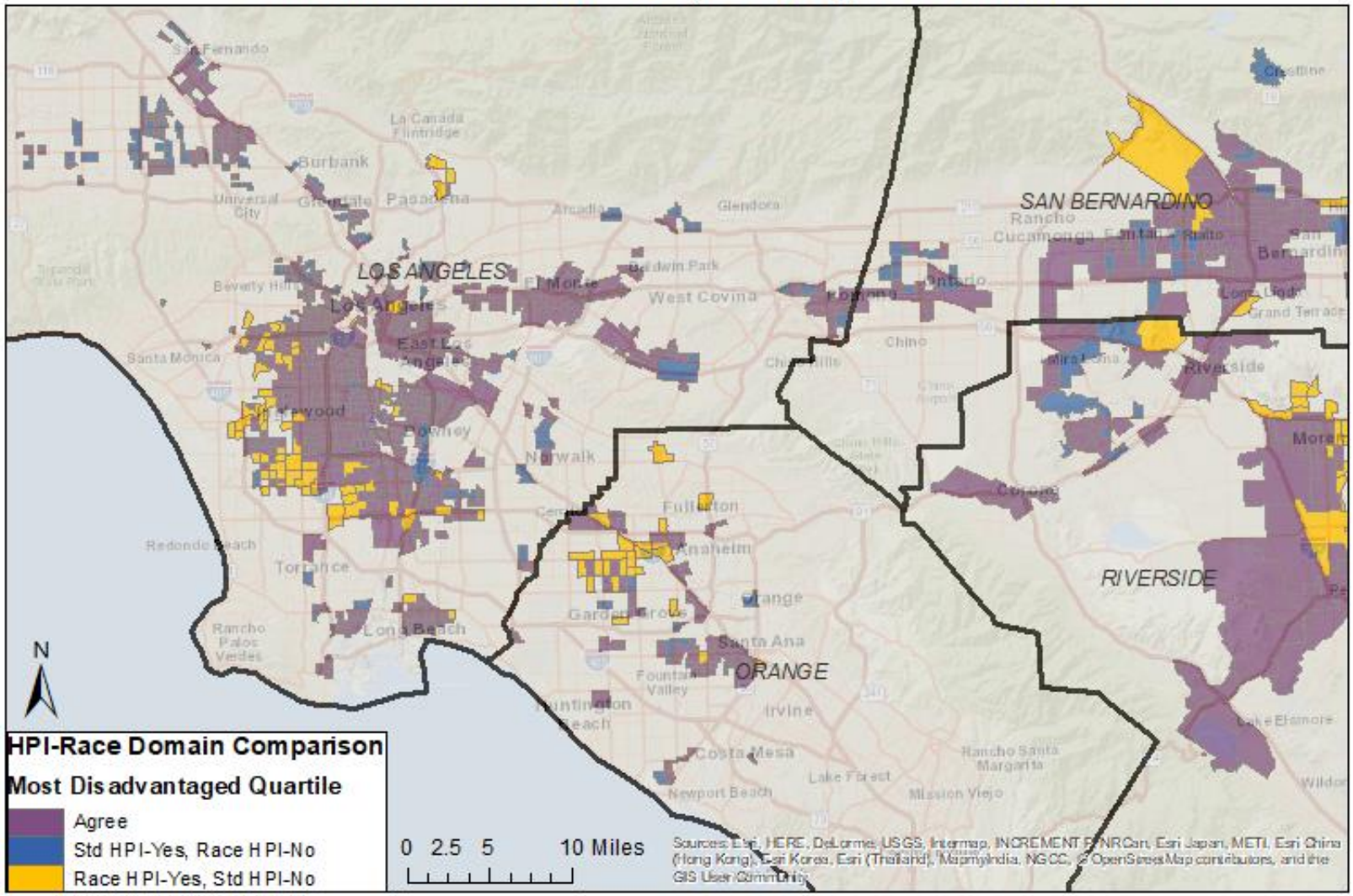
Concordance of Most Disadvantaged Quartile of Census Tracts Standard HPI vs.Race HPI, Lower San Joaquin Valley Counties



public health alliance of southern california

2/6/2018

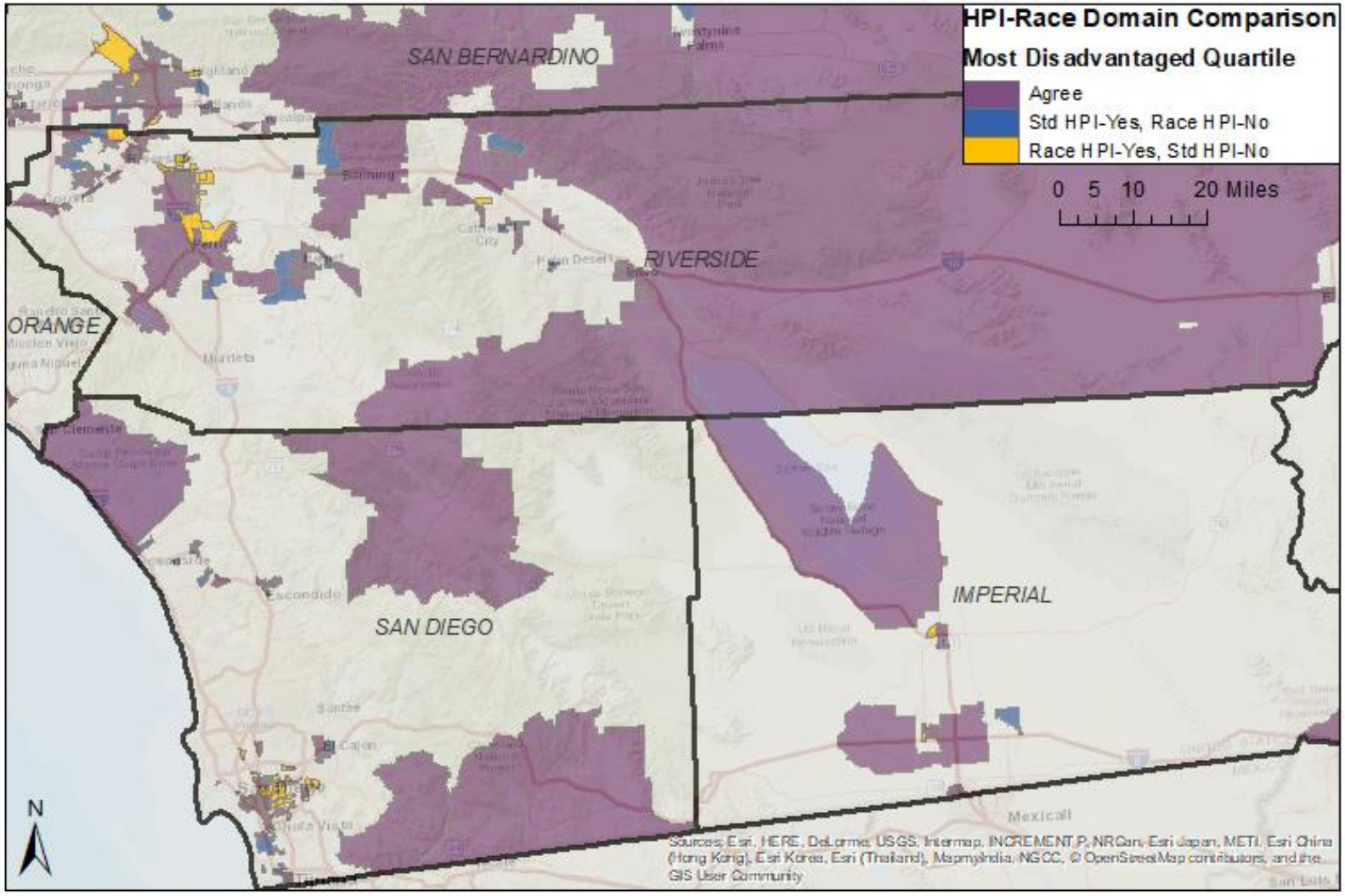
Concordance of Most Disadvantaged Quartile of Census Tracts
Standard HPI vs. Race HPI, Los Angeles Area Counties



public health alliance of southern california

2/6/2018

Concordance of Most Disadvantaged Quartile of Census Tracts
Satndard HPI vs. Race HPI, Southern California



E. HPI Briefs: "The California Healthy Places Index: A New Tool"

The California Healthy Places Index

A New Tool for Community Organizations and Advocates

Every Californian should have the opportunity to live a healthy life. The Healthy Places Index is a powerful new tool to make this a reality, allowing community organizations and advocates to diagnose and change community conditions that affect health and wellbeing. To learn more, visit **HealthyPlacesIndex.org**

How does this tool relate to residents and community advocates?

Community-based organizations and advocates understand that inequities by race or ethnicity, socioeconomic status, and neighborhood are complex and often entrenched. They reflect both current and historic inequities in education, income, food security, housing costs, transportation access, clean air and water, safety net programs, and more. Likewise, all of these conditions shape opportunity for good health and overall wellbeing.

Community residents and coalitions, community development organizations, neighborhood associations, faith-based organizations, service providers, and state and local foundations press local leaders to pursue policies and programs to reduce health inequities and increase opportunity, but which issues and locations should they prioritize? Until now, there have been few convenient tools for evaluating factors at the local level that are known to predict health outcomes, or for gauging how much they matter to community health and wellbeing.

The Healthy Places Index, developed by the Public Health Alliance of Southern California, is a new tool that community organizations and advocates statewide can use to explore local factors that predict life expectancy. HPI was developed based on input from state and local government policy and health equity experts, and provides overall scores as well as detailed data on specific policy action areas that shape health (see figure 1 above). The HPI website offers other resources that advocates and those in a variety of sectors will find useful, including an interactive online map, graphs, data tables, and a policy guide with practical solutions for improving community conditions.

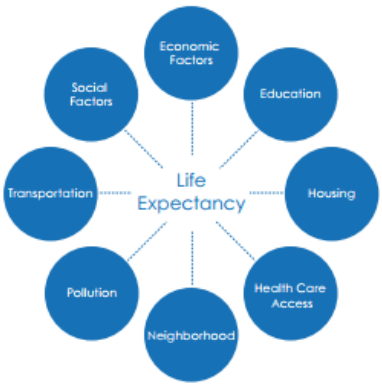


Figure 1: The Healthy Places Index score is derived from 25 variables in eight policy action areas

public health alliance of southern california *A Partnership for Healthy Places*

The California Healthy Places Index

How can this tool be used by community advocates?

Mapping Tool

Identify hotspots where health needs are greatest

HPI Score

Compare conditions with other areas of California and dig deeper to understand community conditions and needs

Move into Action

Policy guide to improve community conditions

Organizations and advocates committed to reducing inequity, promoting opportunity, and improving health—from local residents to large civil rights organizations—can use the HPI in their work empowering communities to understand and improve the conditions that shape health. More specifically, they can use HPI data in conjunction with resident experiences to understand community needs, and to target their energies to the places and conditions where they will have the greatest impact in improving health and equity. For instance, the Healthy, Equitable, Active Land Use Network in Los Angeles has used the HPI in their strategic plan for targeting communty investments.

Community-based organizations can also use the HPI to conduct high-level analysis and produce maps that show health inequities in the communities they serve and use this information to influence policymakers or apply for grants. The multiple levels of data available from a broad set of diciplines can be used to inform the development of campaigns and coalitions. State organizations that span multiple communities or regions of California can use the tool to study the different areas they serve and advocate for increased resources and greater investments in these locations—or suggest that the tool is used to allocate funding and other programs to areas where they will have the greatest impact. For instance, the California Environmental Justice Alliance cited the HPI as a tool for incorporating public health into environmental justice elements of general plans (as required by SB 1000) and for prioritizing communities for programs, policies and funding.

The Healthy Places Index was created in partnership with Virginia Commonwealth University's Center on Society and Health. Learn more about the tool and access the above resources at **HealthyPlacesIndex.org**

public health alliance of southern california *A Partnership for Healthy Places*

F. Recommendations for Future Improvements of the HPI

There are several areas that future versions of the Healthy Places Index (HPI) could explore. First among these is updating the indicators and outcome used in its computation. Calibrating the index based on current life expectancy, or even multiple outcomes, would result in a closer temporal match between indicators and outcome(s), which will likely reduce measurement error. HPI 2.0 was intentionally parsimonious in its indicator selection process in order to produce a policy-relevant index with actionable indicators. Future versions could review available indicators and expand the list used in index computation.

In addition to expanding the list of variables used to compute the index, a second potential area of expansion is to include individually-measured indicators in the index. Currently, all of the indicators in the HPI are Census-tract level aggregates, which were chosen to quantify the health opportunity of a given neighborhood via place-based factors that influence health and well-being. As a result, the HPI 2.0 will provide useful information to guide place-based policies that will improve neighborhoods and, in turn, population health. But measurement occurring only at the tract-level leaves out discussion of potentially important policies that affect individuals such as the earned income tax credit, minimum wage, universal preschool, etc. For example, we know that communities with lower income or lower educational attainment have poorer health. But from a policy perspective, it is important to also know the relative contribution of individual determinants like income and education in the context of various neighborhood factors.

A third area to explore is effects of indicators operating at multiple levels. Many health outcomes are investigated multi-factorially in a *horizontal* sense, with potential causes or influences measured on the same level as the outcome. A relatively more recent trend is to investigate multi-factorial phenomena *vertically* in the sense that causes and influences can arise beyond the individual level (e.g. cohabitation, social determinants of health, environment) at measurement levels different from the outcome. Multilevel modeling is a general framework for combining information from each of several, nested units of measurement (e.g. person, neighborhood, census tract, county, state) in a manner that can allow investigation of associations between measurements at those various levels in a single model. By incorporating more information from different sources, multilevel models can in general improve upon the predictive abilities of models using data from only one level.[63] Multilevel models have been used in many public health research projects, including as examples those from Yang et al.[64]), who investigated predictors of individual survival time at individual, household, parish and municipal levels, and also from Lin et al.[65], who used small area estimation approaches to aggregate individual BRFSS data to estimate county-level quality of life probabilities.

Through including information from various levels, associations and predictions can be aggregated (or de-aggregated) to any particular level. Thus, there is the potential to update the HPI with a combination of both individual and census tract data (or data from any other place-based source) by selecting indicators and deriving domain weights based on importance at the individual level, while maintaining the ability to interpret domain weights at more actionable levels, such as neighborhood, census tract, county, or higher. Individual-level data collected appropriately and at large scale could complement the census-tract information currently included in the HPI to obtain more precise and nuanced relationships between domains and health outcomes, and thus more refined HPI rankings.

Though potential benefits from multilevel models include improved precision and estimates, as well as better prediction, their use is not without cost or limitation. Measurements at finer levels

of measurement (individual, household, neighborhood) are often more difficult to collect across the entire study area, so sampling strategies at these levels would be required. Regardless of the sampling strategy quality, this approach generally leads to more sparse data at the individual data than at higher levels (e.g. census tract, county) and also tends to be costly in terms of collecting measurements. There is the possibility that some individual- or family-level data could be obtained from state-wide or national databases (e.g. BRFSS or the California Health Interview Survey), though there is no guarantee of (i) complete overlap between individual measurements and census-tract indicators, or (ii) that measurements can be linked or geocoded to higher levels.

An additional possibility could be to explore methods to interpolate census tract or county-level indicators onto smaller units of geography (such as neighborhood or block). While such an approach could lead to analysis at more precise and actionable geographic areas, and possibly lead to more nuanced and targeted policy recommendations, it would come at the cost of decreased estimator precision (i.e. more statistical noise). The complicated collection of data management strategies and statistical methods required to conduct this type of analysis would also have limited reproducibility.

A fourth area to explore with both the current and future versions of the HPI is integration with clinical care. Despite an increasing awareness of the importance of integrating primary care and public health [66], most practices have a poor understanding of the communities they serve and their communities' unmet behavioral, social, economic, and environmental needs. Clinicians generally lack the tools to understand the communities that they serve. The HPI could potentially help primary care practices visualize the context in which their patients are living, which could be incorporated into their patient discussions and inform their recommendations. For example, knowledge of air quality could inform discussions with an asthma patient. Future versions of the HPI could provide other indicators (e.g., fast food or grocery store density; crime rates, etc.) that are important to consider when discussing changes to diet and exercise. The HPI could also be used by hospitals as part of risk modeling for patient readmission. Studies have shown that including place-based social determinants improves the predictive power of readmission models using individual patient-level characteristics alone.[67]

References

63. Bryk AS, Raudenbush SW. *Hierarchical Linear Models: Applications and Data Analysis Methods, Second Edition*. Thousand Oaks, CA: Sage Publications; 2002.
64. Yang M, Eldridge S, Merlo J. Multilevel survival analysis of health inequalities in life expectancy. *International Journal for Equity in Health.* 2009;8(31):doi:10.1186/1475-9276-8-31.
65. Lin YH, McLain AC, Probst JC, Bennett KJ, Qureshi ZP, Eberth JM. Health-related quality of life among adults 65 years and older in the United States, 2011-2012: a multilevel small area estimation approach. *Annals of Epidemiology.* 2017;27:52-57.
66. Gofin J, Gofin R, Stimpson JP. Community-oriented primary care (COPC) and the Affordable Care Act: An opportunity to meet the demands of an evolving health care system. *Journal of Primary Care & Community Health.* 2015;6(2):128–133.
67. Hu J, Kind AJH, Nerenz D. Area Deprivation Index predicts readmission risk at an urban teaching hospita. *American Journal of Medical Quality.* 2018; http://journals.sagepub.com/doi/abs/10.1177/1062860617753063.