Curation of new green space indicator for the U.S.: Accessible & recreational park cover (PAD-US-AR)

Matthew H. E. M. Browning[1,2*], Alessandro Rigolon[3], Scott Ogletree[4], Ruoyu Wang[5], Jochem O. Klompmaker[6,7], Chris Bailey[2], Ryan Gagnon[1], Peter James[6,8]

[1] Department of Parks, Recreation and Tourism Management, Clemson University, Clemson, South Carolina 29634, USA
[2] NatureQuant LLC, Bend, OR
[3] Department of City and Metropolitan Planning, The University of Utah, 375 South 1530 East, Salt Lake City, Utah 84112, USA
[4] OPENspace Research Centre, School of Architecture and Landscape Architecture, University of Edinburgh, Edinburgh, UK
[5] UKCRC Centre of Excellence for Public Health/Centre for Public Health, Queen's University Belfast, Belfast, Northern Ireland, United Kingdom
[6] Department of Environmental Health, Harvard T. H. Chan School of Public Health, 655 Huntington Avenue, Boston, Massachusetts 02115, USA
[7] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, 181 Longwood Avenue, Boston, Massachusetts 02115, USA
[8] Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, 401 Park Drive, Boston, Massachusetts 02215, USA
[*] Corresponding author: mhb2@clemson.edu

Abstract

Most spatial epidemiological studies of nature-health relationships use generalized green space measures. For instance, the normalized difference vegetative index (NDVI) is prominent despite its criticisms, such as its inability to differentiate more public (accessible) vs. private (largely inaccessible) land. Green space's capacity to improve health includes building capacities for health-promoting behaviors (e.g., physical activity). Such behaviors may be best activated by recreational and accessible parks.

We curated the Parks and Protected Areas Database of the U.S. (PAD-US) to identify parks that are accessible for outdoor recreation. Our title adds "AR" to "PAD-US" where A=Accessible and R=Recreational. We validated the PAD-US-AR by comparisons with other datasets and demonstrated its uniqueness from other metrics through correlational analyses.

The PAD-US-AR presents a reliable estimate for exposure to parks accessible for outdoor recreation. It has strong associations with home prices, shares of female residents, and shares of older residents, which should be considered as covariates/confounders. The dataset can be a companion to other green space metrics in environmental epidemiology and allied fields of research.

Background & Summary

Exposure science has historically measured the toxic elements that negatively impact human health (Silva et al., 2018). Yet green space and other natural environments represent exposures that can positively influence human health. Research on the health benefits of green space has grown since the 1990s (Taylor & Hochuli, 2017; Zhang et al., 2020). Hundreds of health outcomes/endpoints have been studied, and at least 40 systematic reviews and meta-analyses have been conducted (Twohig-Bennett & Jones, 2018; Yang et al., 2021).

Despite the growing interest in green space and health, the level of sophistication that many researchers approach green space exposure measurement remains low. One prominent and simple exposure measure is the normalized difference vegetation index (NDVI), using satellite imagery to measure the amount of leafy green vegetation (Holland et al., 2021; Labib et al., 2020; Markevych et al., 2017). The calculation of NDVI involves determining the ratio between near-infrared and red bands of light (Jackson & Huete, 1991). NDVI measures hold some value but are limited in several respects. In defense of NDVI, values have been ground-truthed by environmental psychologists and found to correspond to ratings of "greenness" (Rhew et al., 2011). Values can also be easily obtained from Google Earth Engine (GEE) at different spatial and temporal scales across the globe. In critique of NDVI, values cannot indicate the type of, quality of, access to, and experience with vegetation (Holland et al., 2021; Markevych et al., 2017). These limitations should not be surprising; after all, the calculation of NDVI emerged from agricultural science to estimate crop productivity and expected yield rather than environmental epidemiology (Jackson & Huete, 1991). Also limiting NDVI is its inability to identify design characteristics that activate instorative effects of nature-based recreation, such as physical activity along greenways and social interaction at picnic shelters (Browning et al., 2022; Ekkel & Vries, 2017). NDVI values are affected by complex interactions between other environmental factors with less relevance to nature exposure, such as season, slope, and precipitation (Dzhambov et al., 2020; Kumari et al., 2020) in addition to sensor type and the spatial unit size (Helbich, 2019; Su et al., 2019).

Another measure of green vegetation is remotely sensed tree canopy cover. These data are easily retrieved like NDVI and can measure a specific type of greenery by classifying vegetation over a certain height (e.g., >2m) as a tree. Canopy cover is an appropriate green space metric given its affordances for health promotion through shade and psychological restoration (Orians, 1980; Townsend & Barton, 2018). However, like NDVI, tree canopy cover does not provide information on public access and recreational opportunities.

Other advances in the calculation of green space exposure have been made. For instance, machine learning algorithms have been increasingly applied to 360-degree images along streets (e.g., Google Street View [GSV] or Baidu) to calculate the percentage of visible greenery (He & Li, 2021; Labib, 2021; Shahtahmassebi et al., 2020). Still, most green space measures remain limited to the greenery cover and not public and recreational affordances. The need for alternative datasets remains.

Particularly useful would be nationwide data on the location of accessible green spaces managed for outdoor recreation (i.e., parks and protected areas). While the composition and affordances of parks vary, many are managed explicitly for the affordances that are proposed as possible mechanisms explaining the health benefits of green space, including social interaction and physical activity (Cohen et al., 2016; McCormack et al., 2010; cf. Nieuwenhuijsen et al., 2017). In contrast, green spaces in rural areas may be used for resource extraction or conservation without opportunities for recreation and therefore provide little health benefits (Becker et al., 2022). Similarly, green spaces in urban areas can be used for ecosystem services such as stormwater runoff, cooling, and noise/air pollution mitigation but have not been strongly linked to health (Nieuwenhuijsen, 2020).

Researchers are beginning to use some spatial nationwide datasets for measuring park cover in the U.S. (**Table 1**). USA Parks is developed by the Environmental Systems Research

Institute (ESRI) using proprietary data from that company and TomTom. Open Street Map (OSM) includes crowdsourced data tagged by keys (topic/category) and values (features). These can be selected to identify possible public green spaces. The accuracy and consistency of tags vary geographically and are often imprecise, making the identification of public green spaces difficult (Ludwig, Fendrich, et al., 2021). ParkServe contains data on local parks in nearly 14,000 cities, towns, and communities in the USA and was curated by the Trust for Public Land (TPL). Finally, the Parks and Protected Areas Database United States (PAD-US) is an initiative of the U.S. Geological Society (USGS) with federal, state, and local partners. It hopes to inventory all protected areas, including public lands, and voluntarily provide private protected areas.

These currently available park datasets are limited in their ability to identify where accessible and recreational green space exists. Most lack metadata on whether each land parcel is open to the public. OSM provides some data on public access but without clear assignments. For example, our retrieval of polygons with the "leisure:park" tag returned 17 types of access from "community" and "discouraged" to "permissive," "yes," "restricted," and "unknown." Further, OSM data are crowdsourced and not validated by the agencies who manage these spaces. ParkServe also has public access metadata, but its coverage is focused on municipalities. Park cover in rural areas where many important recreational parks (i.e., National Parks) are located is limited in ParkServe.

In response to the value of park data and limitations with extant datasets, we present a new green space indicator – the Protected Areas Dataset US Accessible and Recreational (PAD-US-AR) – for the continental United States. This dataset provides the location of green spaces accessible for recreational purposes. We validate it by comparing it to its source dataset (the original PAD-US), other green space metrics, including NDVI, tree canopy cover, and alternative park datasets, and sociodemographic characteristics in counties and states across the continental U.S.

Table 1. Description of park cover datasets for in the continental U.S.

Name	Developers	Updated	Description	Source	License	URL
USA Parks	Environmental Systems Research Institute (ESRI)	09-2021	"National and State parks and forests, along with County, Regional and Local parks within the United States... provides thousands of named parks and forests at many levels."	ESRI, TomTom	Esri Master License Agreement	https://www.arcgis.com/home/item.html?id=578968f975774d3fab79fe56c8c90941
OSM	Open Street Map (OSM)	06-2022	Park data are available by selecting relevant tags, which consist of a key and value that are separated by a colon. The key is a topic, category, or type of feature (i.e., areas used for leisure). The value provides detail for the key-specified feature (i.e., park vs. playground, both of which are used for leisure). Tags used in past research on park cover and green space measures vary but can include leisure=park, leisure=garden, landuse=grass (Ludwig, Fendrich, et al., 2021); landuse=village_green, and landuse=cemetery (Ludwig, Hecht, et al., 2021); playground and protected_area (Venter et al., 2022); dog park and flower bed (Kraemer & Kabisch, 2021); and allotment, farmland/farmyard, forest/wood, greenfield, greenhouse, meadow, nature reserve, orchard, plant nursery, scrub, village green, and wetland (Zhou et al., 2021). Golf courses have been excluded from some green space analyses (Williams et al., 2020).	Crowdsourced	Open Database License	https://www.openstreetmap.org
ParkServe	Trust for Public Land (TPL)	06-2022	"a comprehensive database of local parks in nearly 14,000 cities, towns, and communities… attempted to contact each city, town, and community with a request for their parks data. If no GIS data was provided, [TPL] created GIS data for the place based on available resources, such as park information from municipal websites, GIS data available from counties and states, and satellite imagery."	Municipal, county, and state GIS datasets; Satellite imagery	Copyright held by the TPL; Data available for personal, non-commercial use	https://www.tpl.org/parkserve/downloads
PAD-US V2.1	United States Geological Survey (USGS)	09-2020	"Nation's inventory of protected areas, including public land and voluntarily provided private protected areas… an ongoing project with several published versions of a spatial database including areas dedicated to the preservation of biological diversity, and other natural (including extraction), recreational, or cultural uses, managed for these purposes through legal or other effective means… its scope expanded in recent years to include all public and nonprofit lands and waters… strives to be a complete inventory of public land and other protected areas, compiling 'best available' data provided by managing agencies and organizations."	Federal, state, and local agencies; National Conservation Easement Database; ParkServe	Public domain	https://www.usgs.gov/programs/gap-analysis-project/science/pad-us-data-download
PADUS-AR V1	The Authors and Curated USGS Data	12-2020	A curated version of the PAD-US to identify parks open to the general public (albeit some with fees for use) and managed for recreational use.	PAD-US V2.1	Creative Commons Attribution 4.0 International	https://osf.io/pwdsg/

Notes: Descriptions were retrieved on June 6, 2022.

Methods

The PAD-US-AR dataset was curated from the USGS Protected Areas Database of the U.S. V2.1 (PAD-US) (U.S. Geological Survey, 2020). These geographic information system (GIS) spatial data compiles the best available data provided by U.S.-based land management agencies and organizations and strives to be a complete inventory of public land and other protected areas. Critically, it includes a field as to whether each park or protected area is publicly accessible, requires a permit to access, or has unknown public access. The V2.1 release became available in September 2020 and included notable updates from previous versions. These included integration of the TPL ParkServe dataset, Census American Indian/Alaskan Native Areas, Ducks Unlimited protected areas, and federal land ownership updates, among others. In contrast, V3.0 was released in early July 2022 and contained only minor updates that we expected to influence our curation process very little. For a full description of version updates, see https://www.usgs.gov/programs/gap-analysis-project/pad-us-data-history.

The PAD-US has been used for conservation mapping (Belote et al., 2016; Dietz et al., 2015; Martinuzzi et al., 2015; Ogletree et al., 2019; Sohl et al., 2014; Theobald, 2014; Walls et al., 2020) and noise research (Buxton et al., 2017; Rice et al., 2020). We are also aware of green space-health studies that have utilized the complete PAD-US dataset (Tsai et al., 2019, 2021). In these studies, the authors identified park locations and ground-truthed results with Google Maps and county/municipal data to identify park entrances.

The opportunities and lack of precedent for curations of the PAD-US prompted us to define which types of parks and protected areas in the dataset were both accessible and recreation-oriented. Based on discussions among three authors (M.B., A.R., S.O.) and four outdoor recreation specialists in the western United States, we reached a consensus on including the following categories:

1. Parks and open spaces open for public access or restricted access (i.e., seasonally open, fees required, or permits required), including but not limited to lands managed by the National Park Service, U.S. Forest Service, Bureau of Land Management, U.S. Fish & Wildlife, Army Corps of Engineers, State Parks, State Departments of Conservation, State Departments of Natural Resources, State Departments of Land, State Fish and Wildlife Departments, State Forest Service, State Park and Recreation Departments, Tennessee Valley Authority, and city and county park and recreation departments.
2. Publicly accessible conservation easements.

We excluded the following designations (see the paragraphs below for rationales):

1. Department of Energy, Department of Defense, and Bureau of Reclamation lands
2. Marine areas are managed by Marine Protected Areas, National Oceanic and Atmospheric Administration, Bureau of Ocean Energy Management, etc.
3. Proclamation areas, which are boundaries of national lands used for administrative purposes that overlap with large areas of public lands that are not all available to the public
4. Fish hatcheries and other lands used for water rights with regulated hunting
5. National Park easements (i.e., lands paralleling but not including the Appalachian Trail and not used by the public)
6. Joint management areas (i.e., university research stations)

7. Non-governmental organization lands (aside from conservation easements)
8. State trust/land survey lands
9. American Indian Lands
10. Other areas with unknown access or closed public access (i.e., limited to coordinated programs and research)

Restricting the PAD-US to these categories was a sequential process starting with the four terrestrial PAD-US domains (**Figure 1**). These domains included designations (policy-designated areas such as National Parks and State Parks), easements (conservation and open space easements provided by the National Conservation Easement Database (*National Conservation Easement Database*, 2022)), fee lands (open space owned by Federal, State, or local agencies, nonprofits, or private individuals), and proclamations (boundaries of administrative areas). For further information on these domains, see http://www.protectedlands.net/pad-us-technical-how-tos/.

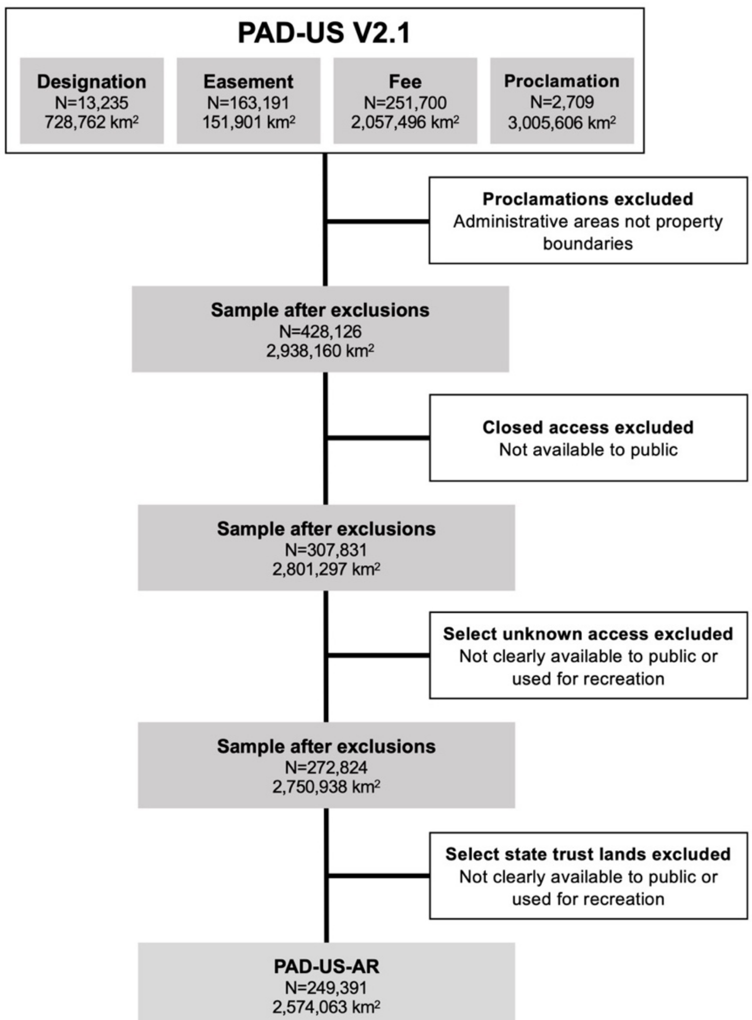


Figure 1. Data curation of the PAD-US-AR to the PAD-US.

Our first step was to exclude all proclamation lands. These administrative boundaries are not ownership lines but are used for agency administrative purposes. Some commercial mapping providers incorrectly use these boundaries to show protected areas and, in doing so, often show large areas of private lands as part of public lands.

Next, we excluded lands described as closed access to public access. Alternative classifications are open to public access or restricted, which denotes a permit is needed or unknown. We temporarily retained unknown access areas for further consideration since large areas of the intermountain west are designated as such. For example, the Great Salt Lake, UT, is the state's largest body water and a recreation destination for boating, swimming, and sunbathing.

The subsequent step was refining such lands labelled as unknown access. Decisions were made based on the assigned land manager. City lands (Code=CITY) were included since many greenways were under this classification. County lands (CNTY), which described nearly 250 polygons run by the City of New York for parks and recreation in the city and upstate, were included. Similarly, regional agency land (REG) covered over 400 polygons that were concentrated in Chicago and Los Angeles suburbs used for parks and recreation; these lands were retained in the dataset. State Department of Conservation (SDC) and State Department of Natural Resource (SDNR) lands were included. These included over 5,000 polygons across the country, including the Great Swamp Management Area, RI, which is an important area for birding and open to the public, and the Great Salt Lake. State Department of Land (SDOL) areas were also included, as they included approximately 30 polygons in Northwestern states used by the public for hiking. State Fish and Wildlife (SFW) lands included urban areas with trails along waterways and were included. State Parks and Recreation (SPR) lands were included and covered public recreational areas in Maine. Tennessee Valley Authority (TVA) and Army Corps of Engineers (USACE) areas covered large reservoirs with important water-based recreation resources and were included. Last, U.S. Forest Service (USFS) lands were retained as they included several recreational areas in Virginia.

All other areas with unknown public access were not deemed accessible to the public and/or used for public recreation and therefore excluded. This conservative approach reduced the chances of misclassification of large tracts of land that likely were inaccessible. For example, Department of Defense (DOD) lands included ammunition plants, Department of Energy (DOE) lands included the nuclear test sites, and National Oceanic and Atmospheric Administration (NOAA) lands that were estuarine research reserves. Non-governmental organization (NGO) lands included nearly 17,500 polygons in the Rocky Mountains but covered too many conservation types to determine whether these were open to the public. American Indian Lands (TRIB) were on reservations and could not be assumed to be accessible and used by the general public.

The final step in curating the PAD-US-AR dataset was determining how to approach the polygons in the Western and Midwestern states that were leftover from the Public Land Survey System (designation = SRMA). Most of these lands follow a grid pattern and are not used for outdoor recreation. However, some state trust lands include important parks, such as DuPont State Forest, NC, which is a popular destination for mountain biking, hiking, swimming, and visiting waterfalls. We manually examined these lands in each state and selected which to include or exclude. Based on this examination, we removed state trust lands from Mississippi, Oklahoma, North Dakota, South Dakota, Montana, Idaho, Wyoming, Utah, Washington, Oregon, Arizona, New Mexico, Texas, Colorado, Louisiana, and Nevada.

To obtain census tract and county exposure estimates, we calculated the percentage of the PAD-US-AR covering each geographic unit. Tract-level estimates included a 0.5-mile buffer around each tract to acknowledge the opportunities for park access for residents living on the tract boundary (Browning & Rigolon, 2019; Rigolon, 2017; Wolch et al., 2013).

Data Records

The data sets are released under the Creative Commons Attribution 4.0 International (CC BY 4.0) license and publicly available. Several files are available:

1. Shapefile and ESRI geopackage
2. CSV of park cover in U.S. counties
3. CSV of park cover in U.S. zip codes
4. CSV of park cover in U.S. tracts with 0.5-mile buffers around each tract

The shapefile includes the original metadata from the PAD-US. For a complete listing, please visit https://www.usgs.gov/programs/gap-analysis-project/pad-us-data-manual. In brief, the data include the name of the parcel; feature class (in the PAD-US-AR, the options are designation, easement, or fee); type and name of management agency (i.e., federal, state, American Indian Lands, or local government); designation (i.e., conversation easement vs. National Park); conservation protection level as designated by the International Union for the Conservation of Nature (IUCN); state name; and geographic size.

The spreadsheets include geographic identifiers (i.e., FIPS codes or GEOID) and percent park cover. Park cover ranges from 0 (no parks) to 1 (complete park cover). Tract estimates are provided for park cover within the boundaries of each tract and the 0.5-mile buffered tract boundaries.

Technical Validation

The PAD-US-AR dataset presents park cover from nearly 250,000 spatial units and 1,900,000 km^2 in area across the continental U.S (**Table 2**). Histograms of the data within counties and tracts and by census region are presented in **Figure S1**. Distributions were right skewed in all regions except Northeastern and Western counties. Northeastern counties showed a flat distribution until approximately 20% cover. Higher levels of cover were present in few counties. Western counties showed an approximately flat distribution until approximately 80% cover, after which the number of counties with higher cover levels was small.

Comparisons with the source dataset are available for each census region in **Figures 2-5.** Large areas of Maine, southeast Pennsylvania, central/western Massachusetts, and northern New Hampshire were excluded from the PAD-US-AR because they were private conservation easements, watersheds with closed access as listed in the PAD-US, or otherwise unknown public access. Swaths of the Dakotas were removed as conservation easements used for wildlife management with uncertain public access. Lands in Oklahoma arranged on a gridwork were removed as state school lands typically leased out for agriculture and mineral resource purposes. A gridwork of land parcels in Montana, Wyoming, Colorado, Arizona, and New Mexico were also removed as state trust lands managed for timber, surface, and mineral resource extraction. Similarly, larger parcels of state trust lands in Western Texas were excluded. Other large parcels of lands excluded were over 560,000 acres in central Idaho, 860,00 acres in southern Nevada, and nearly 200,000 acres in southern South Carolina managed by the Department of Energy;

approximately 550,00 acres at Vermejo Park Ranch managed by Ted Turner Reserves, Inc., and 133,000 areas of the Stronghold District of Badlands National Park in western South Dakota owned by the Oglala Sioux Tribe under agreement by the National Park Service.

We compare the dataset through comparisons with other park datasets, green space metrics, and sociodemographic characteristics. The value of comparing with other park datasets was to determine whether the PAD-US-AR differed from already available datasets. Park dataset comparisons were made by tallying the number of geographic polygon units and calculating the total cover after dissolving all polygon units (to account for some polygons overlapping each other) in census regions.

The value of comparing to other green space metrics was to evaluate whether park cover presented differently than these other common estimates of possible nature exposure. We employed two measures of NDVI (annual averages and summertime highs) and tree canopy cover, which were derived from raster images and averaged across geographic units (tracts or counties). Specifically, values were retrieved and processed in Google Earth Engine (GEE) using cumulative annuals or summertime highs (June-August) from 250x250m 16-day MODIS images averaged over five years (2015-2020) after extracted cloud cover and water pixels. Tree canopy cover was retrieved from the National Land Cover Database (NLCD) 2019 release, which provided estimates for 2016. To identify whether the PAD-US-AR was unique from these other exposure estimates, we examined bivariate correlations between each metric and the PAD-US-AR.

Last, we examined sociodemographic correlates of park cover measured through the PAD-US-AR to inform what confounding factors should be considered when modeling associations between park cover and human health. Sociodemographic characteristics were 2015-2019 American Community Survey (ACS) estimates from the U.S. Census at the county and tract-level. We selected 14 variables based on existing literature examining correlates of green space, especially in studies focused on socioeconomic and racial disparities in access to these spaces (Li et al., 2016; Nesbitt et al., 2019; Pham et al., 2012; Rigolon, 2016; Williams et al., 2020). Attempts at incorporating median household income alongside other measures resulted in multicollinearity so this variable was excluded.

We examined the results of linear mixed models with state as a random effect to account for the hierarchical nature of the data (counties and tracts within states). In all analyses, tracts with fewer than 500 people were removed according to past studies (Los Angeles County, 2022; Thorman & Bohn, 2021) and because small census units provide unreliable estimates of resident's demographic characteristics (Wright & Irimata, 2021). Stratified analyses using urban counties (≥1,000 people/km2) and tracts (≥3000 people/km2) were conducted to inform research in urban areas using thresholds from past research (Browning et al., 2022; Larson et al., 2021).

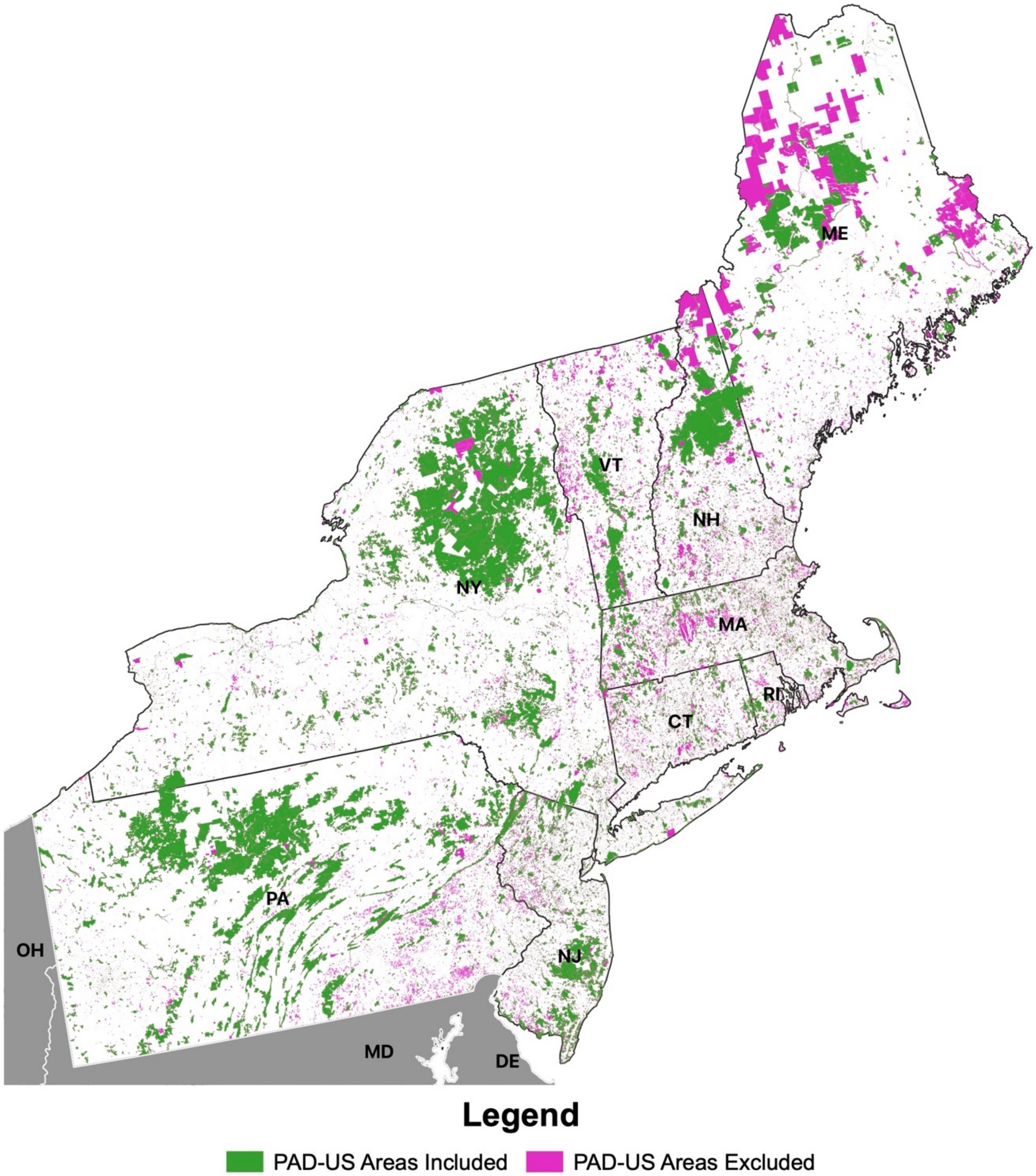


Figure 2. PAD-US-AR park cover dataset compared with its source dataset (PAD-US) in the Northeastern states.

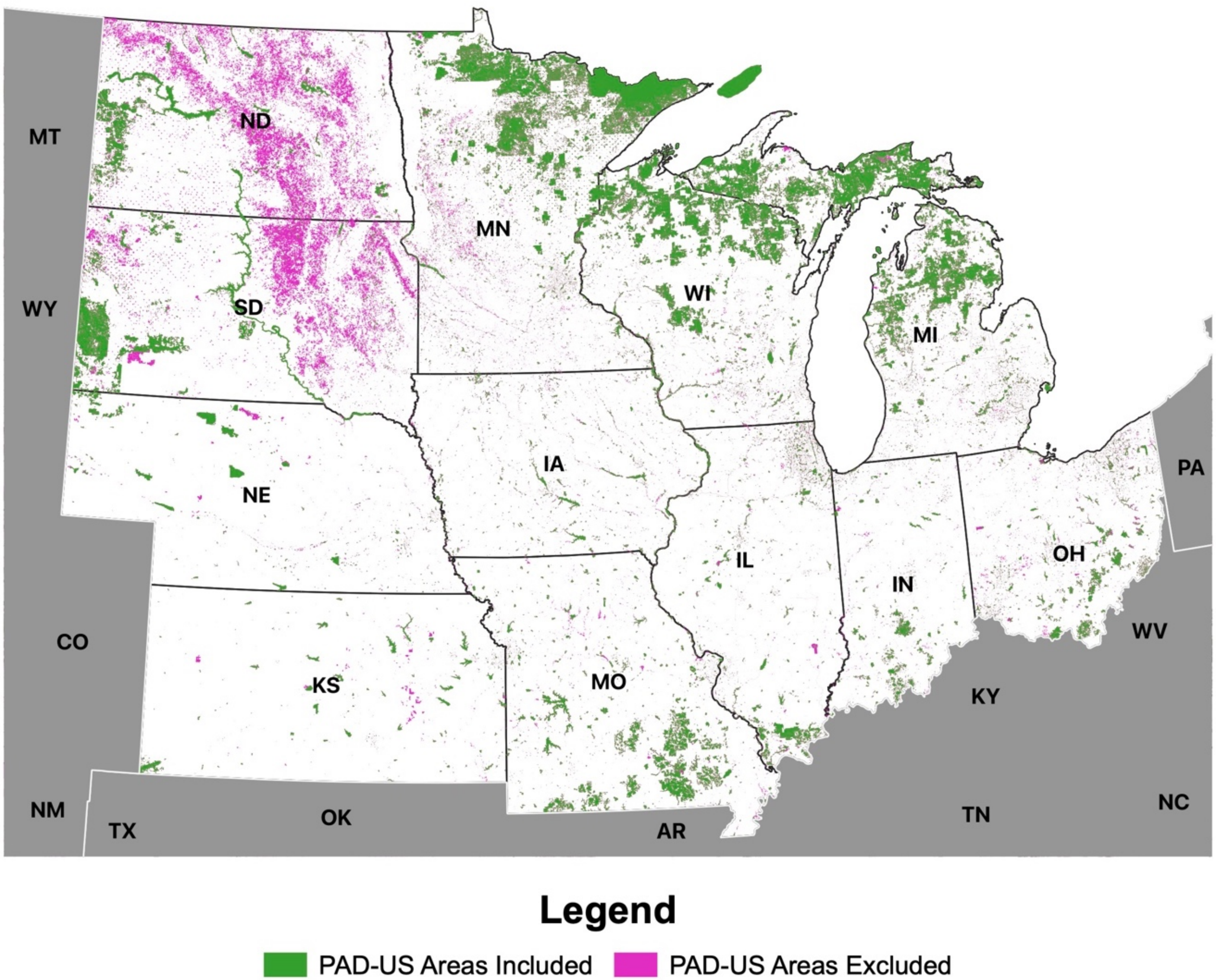


Figure 3. PAD-US-AR park cover dataset compared with its source dataset (PAD-US) in the Midwestern states.

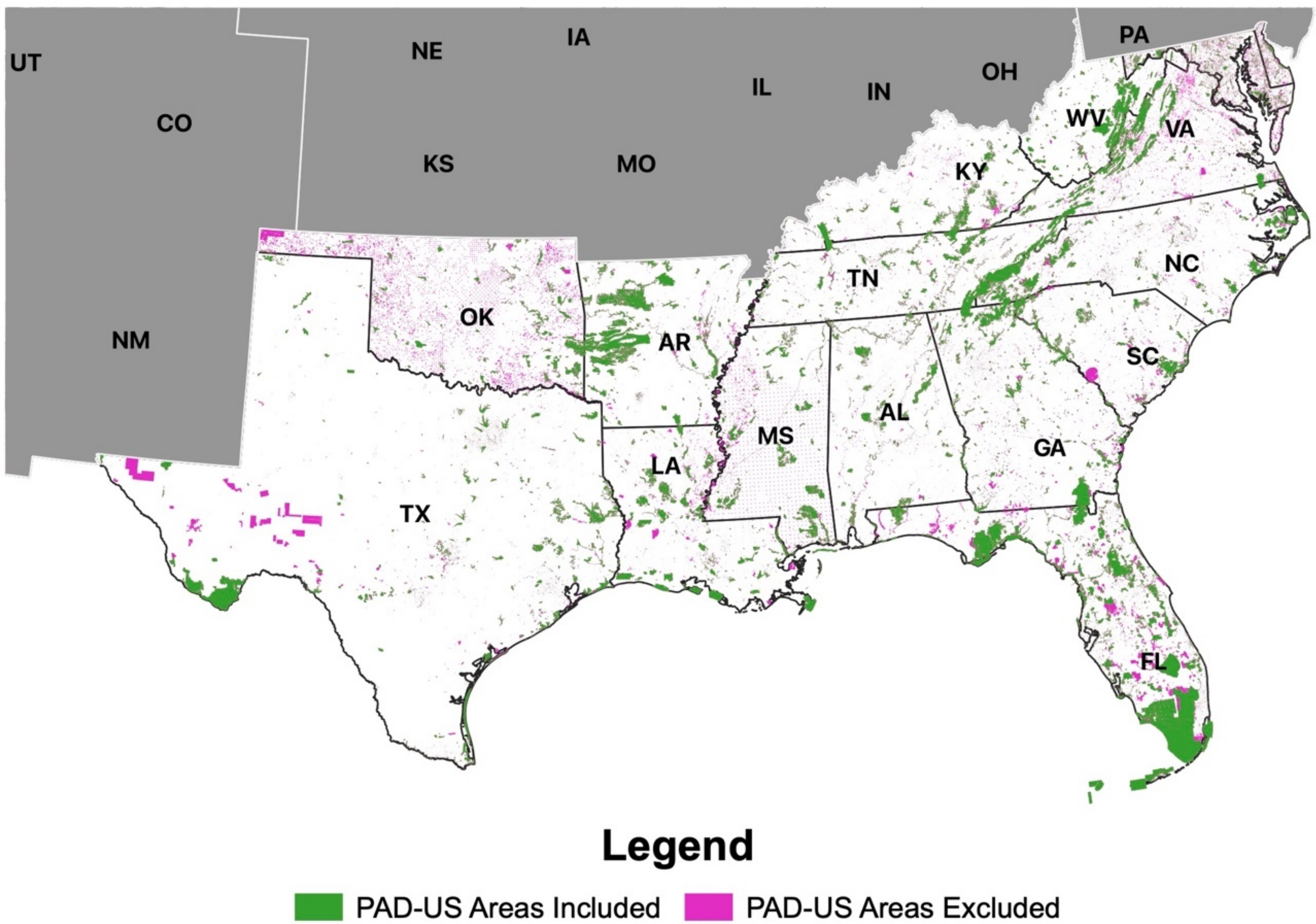


Figure 4. PAD-US-AR park cover dataset compared with its source dataset (PAD-US) in the Southern U.S.

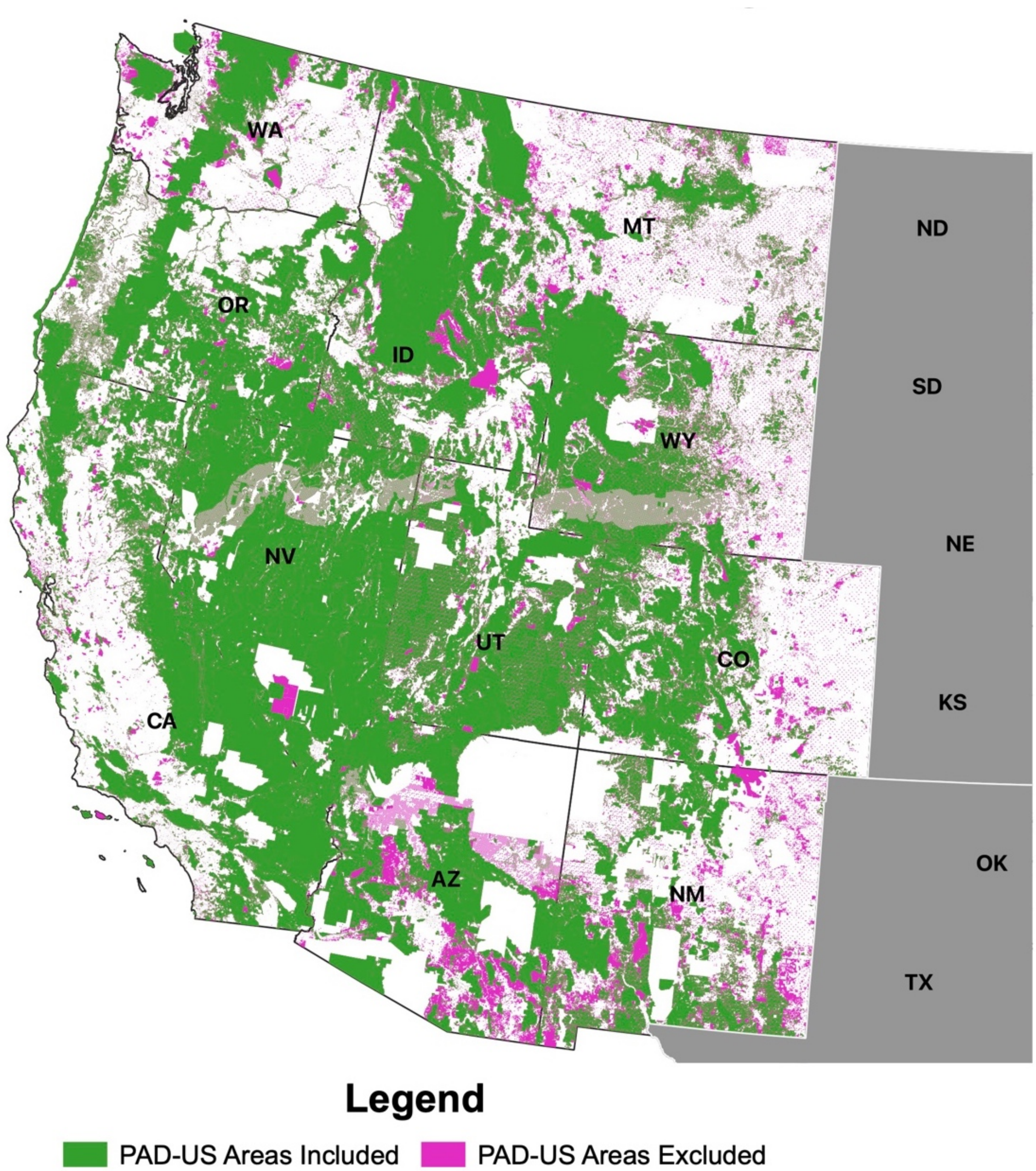


Figure 5. PAD-US-AR park cover dataset compared with its source dataset (PAD-US) in the Western U.S.

Comparison of the PAD-US-AR to other park datasets

Descriptive statistics for each park dataset are provided in **Table 2,** and maps of park cover are provided in **Figure S2**. The PAD-US-AR covers 58.3% of the acreage in the original PAD-US V2.1 dataset. The PAD-US-AR acreage is larger than the acreage of USA Parks and ParkServe but smaller than the OSM datasets when leisure and boundary tags are combined. Bureau of Land Management (BLM) lands are largely absent from the USA Parks and ParkServe datasets but are partially included in the OSM datasets and prominent in the PAD-US-AR. This is particularly noticeable in Nevada, western Utah, and Wyoming. These areas include such popular recreation attractions as the Grand Staircase-Escalante National Monument, UT, and the Grand Canyon Parashant National Monument, AZ. These collectively encompasses nearly 3,000,000 acres (around twice the size of Delaware), attract more than 150,000 visitors annually for hiking, backpacking, and camping, and have received thousands of 5-star reviews on Google Maps. This high number of reviews shows their popularity and visibility in the public sphere. Other notable areas include off-highway vehicle (OHV) trails, such as the Little Sahara OHV Area, UT, which offers driving/riding on a 700-foot drivable sand dune, 30,000 annual visitors, four campgrounds, and approximately 62,000 acres. The majority of popular mountain biking and OHV riding trails around Moab, UT (with the exception of the Slick Rock Trail System) are also BLM lands excluded or with limited coverage from datasets beyond the PAD-US and PAD-US-AR. Collectively, these results demonstrate that the PAD-US-AR presents a selected sample of the PAD-US dataset with differing coverage from pre-existing park cover datasets.

Table 2. Number of units and cover of datasets for park cover in the continental U.S.

Name	Nationwide	Northeast	Midwest	South	West
USA Parks	61,030 (1,049,517 km^2)	9,722 (52,081 km^2)	17,069 (110,524 km^2)	16,639 (172,959 km^2)	17,660 (700,587 km^2)
OSM leisure tags	309,166 (776,436 km^2)	64,442 (51,178 km^2)	90,833 (77,872 km^2)	72,046 (95,613 km^2)	80,113 (522,609 km^2)
OSM boundary tags	51,966 (1,198,021 km^2)	17,174 (69,625 km^2)	8,257 (80,742 km^2)	11,670 (138,321 km^2)	13,760 (890,270 km^2)
ParkServe	135,179 (574,398 km^2)	29,226 (18,672 km^2)	35,246 (52,267 km^2)	37,637 (48,153 km^2)	33,002 (453,441 km^2)
PAD-US V2.1	428,130 (2,211,296 km^2)	110,017 (90,072 km^2)	117,877 (220,224 km^2)	89,880 (226,044 km^2)	109,575 (1,660,575 km^2)
PAD-US-AR V1	249,396 (1,879,299 km^2)	63,464 (68,033 km^2)	69,347 (162,003 km^2)	54,317 (172,760 km^2)	61,643 (1,462,307 km2)

Notes: Number of features/units was derived by selecting by location between census regions and the layer of interest with a negative buffer of 10-m to avoid capturing shared boundaries between region and park features circa June 6, 2022. Some large imprecisely mapped features were manually excluded when they were captured despite negative buffer (min 2 max 6). Areal statistics were calculated by dissolving all polygons and determining cover rather than meta-data provided in the original data. Sum of areas within regions may not total the nationwide statistics in cases of polygons extending beyond the terrestrial area of the U.S. and polygons overlapping regions. OSM tags included dog_park, garden, nature_reserve, and park for the leisure key and national_park and protected_area for the boundary key based on past research utilizing OSM for park cover (Kraemer & Kabisch, 2021; Ludwig, Fendrich, et al., 2021; Ludwig, Hecht, et al., 2021; Venter et al., 2022; Williams et al., 2020; Zhou et al., 2021). PAD-US-AR values differs from Figure 1 because those values were intended to show the number of units/aerial cover lost at each stage of curation while these values were intended to compare park datasets and report results after dissolving park polygons.

Comparison of the PAD-US-AR to other green space measures

Descriptive statistics for park cover in relation to other green space measures are presented in **Table S1,** and maps of each metric are provided in **Figure S3**. Distributions of green space measures are available in **Figures S4-S6**.

Associations between the PAD-US-AR and NDVI varied across geographies and seasons (**Figure 6**). Park cover was negatively associated with NDVI at the county-level (r_{annual} = -.20[-.24, -.17]; r_{summer} = -.33[-.36, -.29]) and not correlated with NDVI at the tract-level (r_{annual} = .03[.02, .04]; r_{summer} = .01[.00, .01]). Associations between the PAD-US-AR and NDVI within census regions were consistently positive, except in Western counties (r_{annual} = -.10[-.19, .00]; r_{summer} = -.00088 [-.11, .09]) or with NDVI summertime maximums in Midwestern counties (r = -.02[-.08, .04]). Such results are likely the result of climatic and land use differences, such as arid climates in the West and high concentrations of agricultural land that only produces chlorophyll in the summer in the Midwest. In juxtaposition, associations between park cover and NDVI annual averages in Midwestern counties were the strongest observed among any pairing (r = .28[.22, .33]). This may be explained by parkland in the upper Midwest having higher concentrations of vegetation that produce chlorophyll year-round (i.e., evergreen trees, wetland herbaceous cover) than in the South and fewer urban parks with less greenery than in the Northeast. Associations at the tract-level ranged from .03[.01, .04] for NDVI summertime maximums in Midwestern tracts to .24[.22, .25] for NDVI summertime maximums in Western tracts.

Park cover was positively associated with tree canopy cover in every pairing. The strongest correlations were among Midwestern counties (r = .65[.61, .68]) and the weakest associations were in nationwide county-level models (r = .10[.07, .14]). The consistent correlation between canopy cover and parks may be explained by people's innate preference for open-growth trees with large amounts of canopy cover (Hofmann et al., 2017; Hull, 1992; Suchocka et al., 2022; Townsend & Barton, 2018) and historic guidelines to retain such trees in park design (Olmsted, 1882).

These findings demonstrate that the PAD-US-AR presents a unique exposure estimate from other green space metrics. Correlations vary in size and direction based on the unit of analysis (counties vs. tracts) and geography (regions of the country and nationwide analyses).

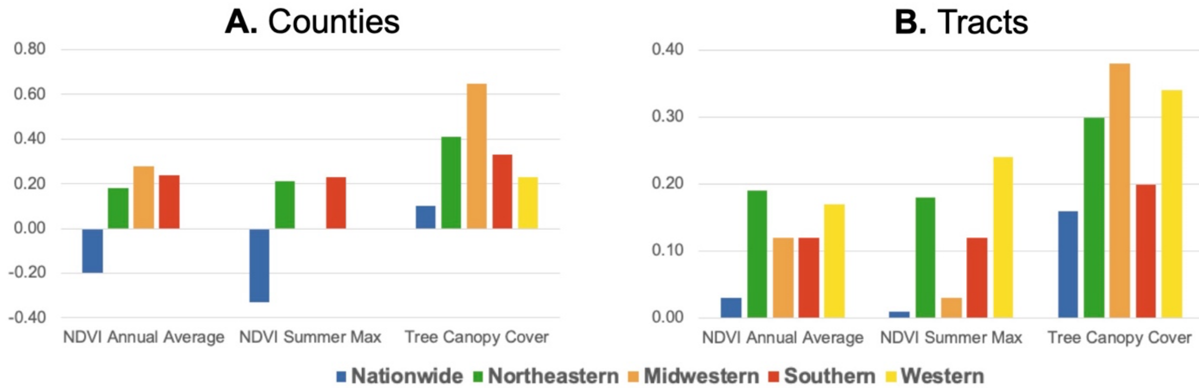


Figure 6. Correlations between park cover and other green space metrics across the continental U.S. within counties (**A**) and tracts (**B**). Notes: Pearson correlation coefficients. Some bars in the figure are not visible when correlations are very small.

Comparison of the PAD-US-AR to sociodemographic characteristics

A listing of the sociodemographic characteristics we examined is provided in **Table S2**. Descriptive statistics for each variable are presented in **Tables S3-S7,** and maps of the distribution of these variables are provided in **Figure S7**.

Multivariate associations between the PAD-US-AR and sociodemographic characteristics are presented in **Figure 7, Figure 8,** and **Table S8**. These results are derived from generalized linear models (GLMs) accounting for state random effects with minimal multicollinearity (**Table S9**).

Park cover was more strongly associated with sociodemographic characteristics at the county-level than at the tract-level. Around 30% of the variance in countywide park cover was explained in U.S. regions ($R^2_{Northeast}$=.29, $R^2_{Midwest}$=.31, R^2_{South}=.23, R^2_{West}=.38). Variance explained within counties across the country was over 60% ($R^2_{Nationwide}$=.63). Variance explained at the tract level was closer to 10%-20% ($R^2_{Nationwide}$=.19, $R^2_{Northeast}$=.09, $R^2_{Midwest}$=.08, R^2_{South}=.12, R^2_{West}=.18).

Three sociodemographic characteristics showed fairly consistent associations with park cover. Areas with greater shares of older adults (≥65 yrs) reliably had more park cover on average. Areas with higher median home values also had more park cover on average, but there were exceptions in the Northeast where home values and park cover was not associated. Last, areas with greater shares of female residents had less park cover on average. The exceptions were Northeastern and Southern counties, where sex and park cover were not associated.

Associations between the PAD-US-AR and other sociodemographic characteristics varied by census region. For example, park cover in Northeastern counties was concentrated in areas with less income inequality, lower shares of high school graduates, and lower shares of people employed in natural resource professions. Park cover in Midwestern counties was greater in areas with higher poverty and unemployment levels. Park cover in Midwestern tracts was concentrated in areas with greater shares of college graduates. Southern counties showed more park cover where poverty levels and population sizes were higher as well as where shares of NH Black residents were lower. Park cover was concentrated in Western counties with greater shares of college graduates and poverty levels and smaller shares of NH Asian residents and income inequality. Park cover in Western tracts was concentrated in areas with lower population densities and smaller shares of Hispanic residents. In summary, park cover is associated with many sociodemographic characteristics, and the strength and direction vary by geography and unit of analysis.

Multivariate associations between the PAD-US-AR and sociodemographic characteristics in urban areas are presented in **Table S10**. Median home value continued to show strong positive associations with park cover in more regions. One exception was observed in Midwestern tracts, where median home value was negatively associated with park cover. Percent female no longer predicted park cover except in Southern tracts. Shares of older adults also predicted park cover in only a few urban cases; significant positive associations were observed only in nationwide and Northeastern tracts. Percent NH Asian residents emerged as a predictor in several models, but the direction of the associations differed. Nationwide models showed a negative association while Midwestern counties and tracts and Southern tracts showed positive associations. County-level models of urban areas continued to predict the variance explained of park cover better than tract-level models of urban areas.

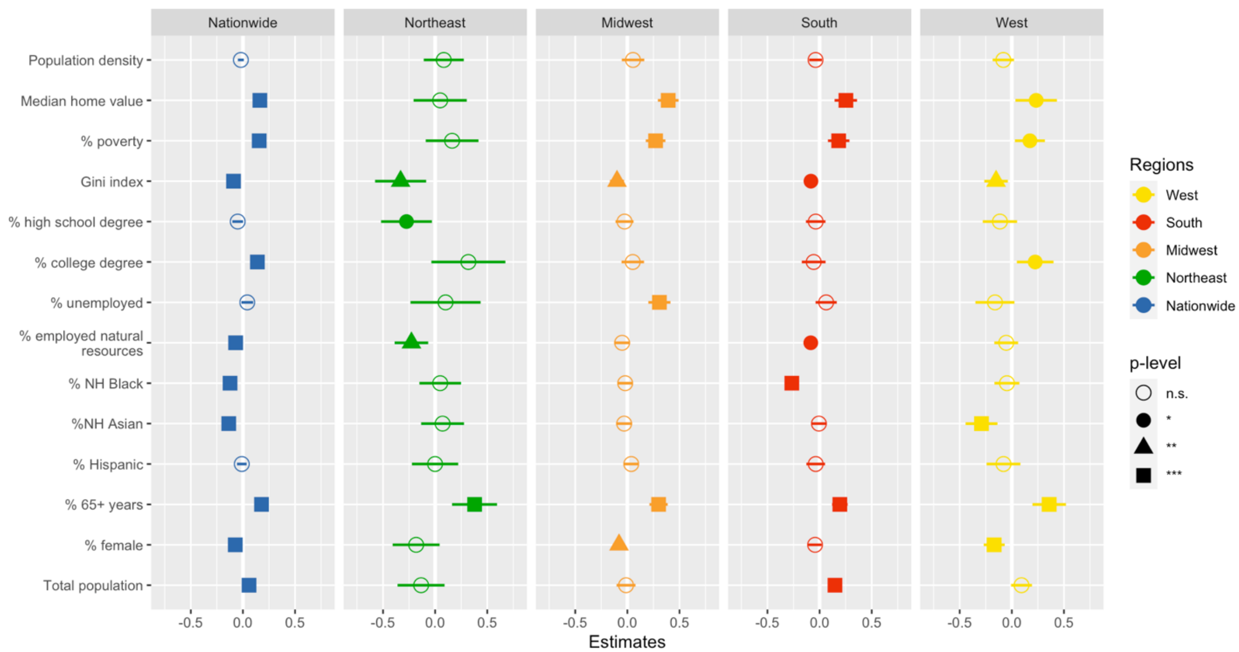


Figure 7. Regressing sociodemographic characteristics on the PAD-US-AR park cover dataset within counties. Notes: Generalized linear mixed models with state random effects. Standardized betas and 95% confidence intervals are shown. Differing symbols represent statistical significance (p-value): empty circle is shown for p >.05, filled-in circle for $p < .05$; triangle for $p < .01$; square for $p < .001$.

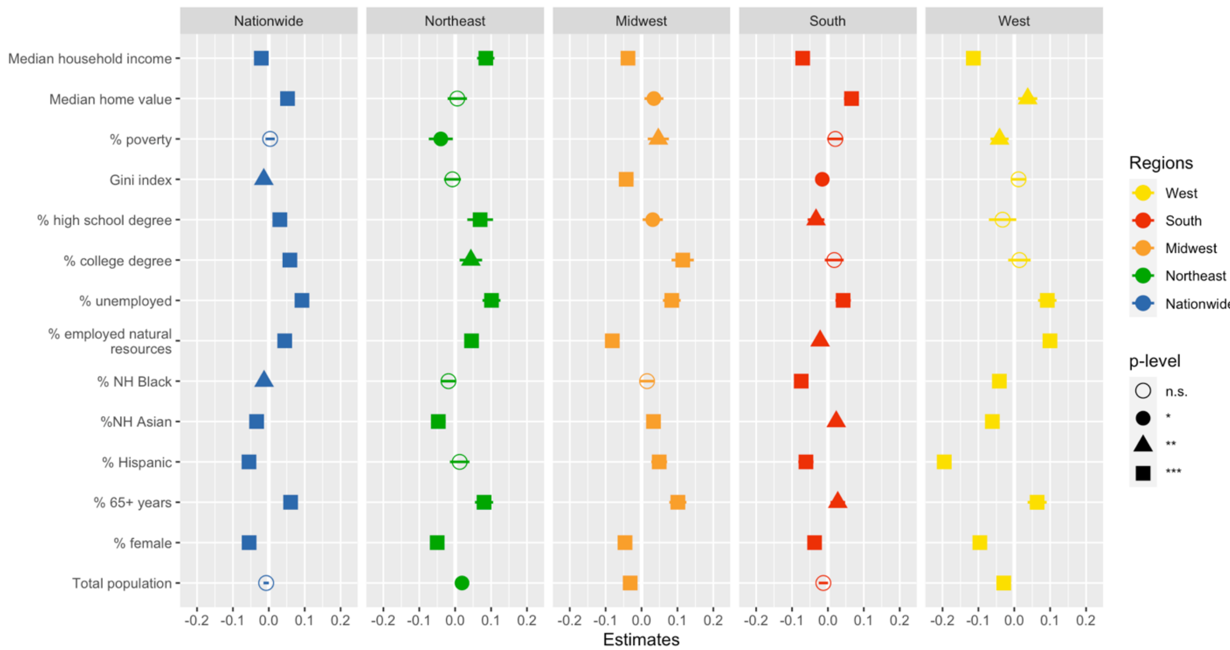


Figure 8. Regressing sociodemographic characteristics on the PAD-US-AR park cover dataset within tracts. Notes: Generalized linear mixed models with state random effects. Standardized betas and 95% confidence intervals are shown. Differing symbols represent statistical significance (p-value): empty circle is shown for $p > .05$, filled-in circle for $p < .05$; triangle for $p < .01$; square for $p < .001$.

Usage Notes

We present a new indicator of green space (PAD-US-AR) for the continental U.S: the location of parks accessible for recreation. This dataset allows researchers to examine not only the quantity of green space around geographic units of interest (homes, neighborhoods, transit routes) but also that these green spaces are designed for and support recreational uses. Other readily-available metrics – like NDVI and tree canopy cover – are unable to identify whether the observed locations of green spaces are usable by the public for health-promoting recreation. The dataset is unique from these other green space metrics, as determined by the bivariate correlations presented above.

The PAD-US-AR also differs in coverage from pre-existing park datasets. The reasons to utilize these data rather than other options include the fact the source data were validated by the agencies managing the land, the systematic examination of what is accessible for recreation, and the clarity and transparency in its curation.

The chances for residual confounding in ecological (area-level studies) might be high if multivariate models do not control for sociodemographic characteristics of the areas encompassing parks. The PAD-US-AR has the strongest and most consistent associations across U.S. regions with home prices, shares of female residents, and shares of older presents. These should be statistically controlled in models including the PAD-US-AR as a covariate. Other measures of socio-economic status (i.e., median household income) might be insufficient to avoid residual confounding in ecological studies.

As the nature/green space and health literature expands, exposure estimates are expected to develop and be refined. The PAD-US-AR presents an important advancement in this body of literature by offering researchers an estimate of where parks are available for outdoor recreation.

Code Availability

The programs used to generate all the results were QGIS (3.18.3), ArcGIS Desktop 10.8.1 and R (4.1.2). Analysis scripts are available on request from M.B. (mhb2@clemson.edu).

Acknowledgments

Clemson University is acknowledged for its generous allotment of computing time on the Palmetto Cluster.

Author Contributions

M.B., A.R., and S.O. conceived the research. M.B., S.O., and C.B. processed the data. M.B. wrote the manuscript. All authors contributed to and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Competing Interests

The authors declare no competing interests.

References

Becker, D. A., Browning, M. H., McAnirlin, O., Yuan, S. & Helbich, M. (2022). Is Green Space Associated with Opioid-Related Mortality? An Ecological Study at the U.S. County Level. *Urban Forestry & Urban Greening*, 127529. https://doi.org/10.1016/j.ufug.2022.127529

Belote, R. T., Dietz, M. S., McRae, B. H., Theobald, D. M., McClure, M. L., Irwin, G. H., McKinley, P. S., Gage, J. A. & Aplet, G. H. (2016). Identifying Corridors among Large Protected Areas in the United States. *PLOS ONE*, *11*(4), e0154223-16. https://doi.org/10.1371/journal.pone.0154223

Browning, M. H. E. M. & Rigolon, A. (2019). Could nature help children rise out of poverty? Green space and future earnings from a cohort in ten U.S. cities. *Environmental Research*, *176*, 1–13. https://doi.org/https://doi.org/10.1016/j.envres.2019.04.016

Browning, M. H. E. M., Rigolon, A., McAnirlin, O. & Yoon, H. (Violet). (2022). Where greenspace matters most: A systematic review of urbanicity, greenspace, and physical health. *Landscape and Urban Planning*, *217*, 104233. https://doi.org/10.1016/j.landurbplan.2021.104233

Buxton, R. T., McKenna, M. F., Mennitt, D., Fristrup, K., Crooks, K., Angeloni, L. & Wittemyer, G. (2017). Noise pollution is pervasive in U.S. protected areas. *Science*, *356*(6337), 531–533. https://doi.org/10.1126/science.aah4783

Cohen, D. A., Han, B., Nagel, C. J., Harnik, P., McKenzie, T. L., Evenson, K. R., Marsh, T., Williamson, S., Vaughan, C. & Katta, S. (2016). The First National Study of Neighborhood Parks. *American Journal of Preventive Medicine*, *51*(4), 419–426. https://doi.org/10.1016/j.amepre.2016.03.021

Dietz, M. S., Belote, R. T., Aplet, G. H. & Aycrigg, J. L. (2015). The world's largest wilderness protection network after 50years: An assessment of ecological system representation in the U.S. National Wilderness Preservation System. *Biological Conservation*, *184*(C), 431–438. https://doi.org/10.1016/j.biocon.2015.02.024

Dzhambov, A. M., Lercher, P., Rüdisser, J., Browning, M. H. E. M. & Markevych, I. (2020). Allergic symptoms in association with naturalness, greenness, and greyness: A cross-sectional study in schoolchildren in the Alps. *Environmental Research*, *198*, 110456. https://doi.org/10.1016/j.envres.2020.110456

Ekkel, E. D. & Vries, S. de. (2017). Nearby green space and human health: Evaluating accessibility metrics. *Landscape and Urban Planning*, *157*, 214–220. https://doi.org/10.1016/j.landurbplan.2016.06.008

He, N. & Li, G. (2021). Urban neighbourhood environment assessment based on street view image processing: A review of research trends. *Environmental Challenges*, 100090.

Helbich, M. (2019). Spatiotemporal Contextual Uncertainties in Green Space Exposure Measures: Exploring a Time Series of the Normalized Difference Vegetation Indices. *International Journal of Environmental Research and Public Health*, *16*(5), 852. https://doi.org/10.3390/ijerph16050852

Hofmann, M., Gerstenberg, T. & Gillner, S. (2017). Predicting tree preferences from visible tree characteristics. *European Journal of Forest Research*, *136*(3), 421–432. https://doi.org/10.1007/s10342-017-1042-7

Holland, I., DeVille, N. V., Browning, M. H. E. M., Buehler, R. M., Hart, J. E., Hipp, J. A., Mitchell, R., Rakow, D. A., Schiff, J. E., White, M. P., Yin, J. & James, P. (2021). Measuring nature contact: A narrative review. *International Journal of Environmental Research and Public Health*, *18*(8), 4092. https://doi.org/10.3390/ijerph18084092

Hull, R. B. (1992). How the public values urban forests. *Journal of Arboriculture*.

Jackson, R. D. & Huete, A. R. (1991). Interpreting vegetation indices. *Preventive Veterinary Medicine*.

Kraemer, R. & Kabisch, N. (2021). Parks in context: advancing citywide spatial quality assessments of urban green spaces using fine-scaled indicators. *Ecology and Society*, *26*(2). https://doi.org/10.5751/es-12485-260245

Kumari, N., Saco, P. M., Rodriguez, J. F., Johnstone, S. A., Srivastava, A., Chun, K. P. & Yetemen, O. (2020). The Grass Is Not Always Greener on the Other Side: Seasonal Reversal of Vegetation Greenness in Aspect-Driven Semiarid Ecosystems. *Geophysical Research Letters*, *47*(15). https://doi.org/10.1029/2020gl088918

Labib, S. M. (2021). Estimating multiple greenspace exposure types and their associations with neighbourhood premature mortality: a socioecological study. *Science of The Total Environment*, 147919. https://doi.org/10.1016/j.scitotenv.2021.147919

Labib, S. M., Lindley, S. & Huck, J. J. (2020). Spatial dimensions of the influence of urban green-blue spaces on human health: A systematic review. *Environmental Research*, *180*, 108869. https://doi.org/10.1016/j.envres.2019.108869

Larson, L. R., Mullenbach, L. E., Browning, M. H. E. M., Rigolon, A., Thomsen, W. A. J., Covelli, M. W. A. E., Reigner, N. P., Sharaievska, I., McAnirlin, O., D'Antonio, A., Cloutier, S., Helbich, M. & Labib, S. M. (2021). Greenspace and park use associated with less emotional distress among college students in the United States during the COVID-19 pandemic. *Environmental Research*, 112367. https://doi.org/10.1016/j.envres.2021.112367

Li, X., Zhang, C., Li, W. & Kuzovkina, Y. A. (2016). Environmental inequities in terms of different types of urban greenery in Hartford, Connecticut. *Urban Forestry & Urban Greening*, *18*, 163–172. https://doi.org/10.1016/j.ufug.2016.06.002

Los Angeles County. (2022). *COVID-19 vaccine dashboard - Census Tract Vaccination Estimates Data File Notes*. http://publichealth.lacounty.gov/media/coronavirus/vaccine/vaccine-dashboard.htm

Ludwig, C., Fendrich, S. & Zipf, A. (2021). Regional variations of context-based association rules in OpenStreetMap. *Transactions in GIS*, *25*(2), 602–621. https://doi.org/10.1111/tgis.12694

Ludwig, C., Hecht, R., Lautenbach, S., Schorcht, M. & Zipf, A. (2021). Mapping Public Urban Green Spaces Based on OpenStreetMap and Sentinel-2 Imagery Using Belief Functions. *ISPRS International Journal of Geo-Information*, *10*(4), 251. https://doi.org/10.3390/ijgi10040251

Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., Vries, S. de, Triguero-Mas, M., Brauer, M., Nieuwenhuijsen, M. J., Lupp, G., Richardson, E. A., Astell-Burt, T., Dimitrova, D. D., Feng, X., Sadeh, M., Standl, M., Heinrich, J. & Fuertes, E. (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental Research*, *158*, 301–317. https://doi.org/10.1016/j.envres.2017.06.028

Martinuzzi, S., Radeloff, V. C., Joppa, L. N., Hamilton, C. M., Helmers, D. P., Plantinga, A. J. & Lewis, D. J. (2015). Scenarios of future land use change around United States' protected areas. *Biological Conservation*, *184*, 446–455. https://doi.org/10.1016/j.biocon.2015.02.015

McCormack, G. R., Rock, M., Toohey, A. M. & Hignell, D. (2010). Characteristics of urban parks associated with park use and physical activity: A review of qualitative research. *Health & Place*, *16*(4), 712–726. https://doi.org/10.1016/j.healthplace.2010.03.003

National Conservation Easement Database. (2022). https://www.conservationeasement.us

Nesbitt, L., Meitner, M. J., Girling, C., Sheppard, S. R. J. & Lu, Y. (2019). Who has access to urban vegetation? A spatial analysis of distributional green equity in 10 US cities. *Landscape and Urban Planning*, *181*, 51–79. https://doi.org/10.1016/j.landurbplan.2018.08.007

Nieuwenhuijsen, M. J. (2020). Green infrastructure and health. *Annual Review of Public Health*, *42*(1), annurev-publhealth-090419-102511-12. https://doi.org/10.1146/annurev-publhealth-090419-102511

Nieuwenhuijsen, M. J., Khreis, H., Triguero-Mas, M., Gascon, M. & Dadvand, P. (2017). Fifty shades of green. *Epidemiology*, *28*(1), 63–71. https://doi.org/10.1097/ede.0000000000000549

Ogletree, S. S., Powell, R. B., Baldwin, R. F. & Leonard, P. B. (2019). A framework for mapping cultural resources in landscape conservation planning. *Conservation Science and Practice*, *1*(6), e41-11. https://doi.org/10.1111/csp2.41

Olmsted, F. L. (1882). Trees in streets and in parks. In *The highest value of a park*.

Orians, G. H. (1980). *Habitat selection: General theory and applications for human behavior* (J. S. Lockard, Ed.; pp. 49–66). Elsevier Science Limited.

Pham, T.-T.-H., Apparicio, P., Séguin, A.-M., Landry, S. & Gagnon, M. (2012). Spatial distribution of vegetation in Montreal: An uneven distribution or environmental inequity? *Landscape and Urban Planning*, *107*(3), 214–224. https://doi.org/10.1016/j.landurbplan.2012.06.002

Rhew, I. C., Stoep, A. V., Kearney, A., Smith, N. L. & Dunbar, M. D. (2011). Validation of the Normalized Difference Vegetation Index as a Measure of Neighborhood Greenness. *Annals of Epidemiology*, *21*(12), 946–952. https://doi.org/10.1016/j.annepidem.2011.09.001

Rice, W. L., Newman, P., Miller, Z. D. & Taff, B. D. (2020). Protected areas and noise abatement: A spatial approach. *Landscape and Urban Planning*, *194*, 103701–103709. https://doi.org/10.1016/j.landurbplan.2019.103701

Rigolon, A. (2016). A complex landscape of inequity in access to urban parks: A literature review. *Landscape and Urban Planning*, *153*, 160–169. https://doi.org/10.1016/j.landurbplan.2016.05.017

Rigolon, A. (2017). Parks and young people: An environmental justice study of park proximity, acreage, and quality in Denver, Colorado. *Landscape and Urban Planning*, *165*, 73–83. https://doi.org/10.1016/j.landurbplan.2017.05.007

Shahtahmassebi, A., Li, C., Fan, Y., Wu, Y., lin, Y., Gan, M., Wang, K., Malik, A. & Blackburn, A. (2020). Remote sensing of urban green spaces: a review. *Urban Forestry & Urban Greening*, 126946. https://doi.org/10.1016/j.ufug.2020.126946

Silva, R. A., Rogers, K. & Buckley, T. J. (2018). Advancing Environmental Epidemiology to Assess the Beneficial Influence of the Natural Environment on Human Health and Well-Being. *Environmental Science & Technology*, *52*(17), 9545–9555. https://doi.org/10.1021/acs.est.8b01781

Sohl, T. L., Sayler, K. L., Bouchard, M. A., Reker, R. R., Friesz, A. M., Bennett, S. L., Sleeter, B. M., Sleeter, R. R., Wilson, T., Soulard, C., Knuppe, M. & Hofwegen, T. V. (2014). Spatially explicit modeling of 1992–2100 land cover and forest stand age for the conterminous United States. *Ecological Applications*, *24*(5), 1015–1036. https://doi.org/10.1890/13-1245.1

Su, J. G., Dadvand, P., Nieuwenhuijsen, M. J., Bartoll, X. & Jerrett, M. (2019). Associations of green space metrics with health and behavior outcomes at different buffer sizes and remote sensing sensor resolutions. *Environment International*, *126*, 162–170. https://doi.org/10.1016/j.envint.2019.02.008

Suchocka, M., Wojnowska-Heciak, M., Błaszczyk, M., Gawłowska, A., Ciemniewska, J., Jarska, A., Heciak, J. & Pachnowska, B. (2022). Old trees are perceived as a valuable element of the municipal forest landscape. *PeerJ*, *10*, e12700. https://doi.org/10.7717/peerj.12700

Taylor, L. & Hochuli, D. F. (2017). Defining greenspace: Multiple uses across multiple disciplines. *Landscape and Urban Planning*, *158*, 25–28. https://doi.org/10.1016/j.landurbplan.2016.09.024

Theobald, D. M. (2014). Development and Applications of a Comprehensive Land Use Classification and Map for the US. *PLoS ONE*, *9*(4), e94628. https://doi.org/10.1371/journal.pone.0094628

Thorman, T. & Bohn, S. (2021). *2020 Census Maps - California's Hard-to-Count Communities Technical Appendix*. Public Policy Institute of California.

Townsend, J. B. & Barton, S. (2018). The Impact of Ancient Tree Form on Modern Landscape Preferences. *Urban Forestry & Urban Greening*, *34*, 205–216. https://doi.org/10.1016/j.ufug.2018.06.004

Tsai, W.-L., Davis, A. J. S. & Jackson, L. E. (2019). Associations between types of greenery along neighborhood roads and weight status in different climates. *Urban Forestry & Urban Greening*, 1–44.

Tsai, W.-L., Nash, M. S., Rosenbaum, D. J., Prince, S. E., D'Aloisio, A. A., Neale, A. C., Sandler, D. P., Buckley, T. J. & Jackson, L. E. (2021). Types and spatial contexts of neighborhood greenery matter in associations with weight status in women across 28 U.S. communities. *Environmental Research*, 111327.

Twohig-Bennett, C. & Jones, A. (2018). The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environmental Research*, *166*, 628–637. https://doi.org/10.1016/j.envres.2018.06.030

U.S. Geological Survey. (2020, December 20). *Protected Areas Database of the United States (PAD-US) 2.1*. https://www.usgs.gov/core-science-systems/science-analytics-and-synthesis/gap/science/pad-us-data-download?qt-science_center_objects=0#qt-science_center_objects

Venter, Z. S., Shackleton, C., Faull, A., Lancaster, L., Breetzke, G. & Edelstein, I. (2022). Is green space associated with reduced crime? A national-scale study from the Global South. *Science of The Total Environment*, *825*, 154005. https://doi.org/10.1016/j.scitotenv.2022.154005

Walls, M., Lee, P. & Ashenfarb, M. (2020). National monuments and economic growth in the American West. *Science Advances*, *6*(12), eaay8523. https://doi.org/10.1126/sciadv.aay8523

Williams, T. G., Logan, T. M. & Zuo, C. T. (2020). Parks and safety: a comparative study of green space access and inequity in five US cities. *Landscape and Urban Planning*.

Wolch, J., Wilson, J. P. & Fehrenbach, J. (2013). Parks and Park Funding in Los Angeles: An Equity-Mapping Analysis. *Urban Geography*, *26*(1), 4–35. https://doi.org/10.2747/0272-3638.26.1.4

Wright, T. & Irimata, K. (2021). *Empirical Study of Two Aspects of the Topdown Algorithm Output for Redistricting: Reliability & Variability (August 5, 2021 updated)*. Center for Statistical Research & Methodology, Research and Methodology Directorate, U.S. Census Bureau.

Yang, B.-Y., Zhao, T., Hu, L.-X., Browning, M. H. E. M., Heinrich, J., Dharmage, S. C., Jalaludin, B., Knibbs, L. D., Liu, X.-X., Luo, Y.-N., James, P., Li, S., Huang, W.-Z., Chen, G., Zeng, X.-W., Hu, L.-W., Yu, Y. & Dong, G.-H. (2021). Greenspace and human health: An umbrella review. *The Innovation*, *2*(4), 100164. https://doi.org/10.1016/j.xinn.2021.100164

Zhang, J., Yu, Z., Zhao, B., Sun, R. & Vejre, H. (2020). Links between green space and public health: A bibliometric review of global research trends and future prospects from 1901 to 2019. *Environmental Research Letters*, *15*(6), 063001–063037. https://doi.org/10.1088/1748-9326/ab7f64

Zhou, Y., Lengerke, T. von & Dreier, M. (2021). Comparing different data sources by examining the associations between surrounding greenspace and children's weight status. *International Journal of Health Geographics*, *20*(1), 24. https://doi.org/10.1186/s12942-021-00278-w

SUPPLEMENTAL RESOURCES

Curation of a new green space indicator for the continental United States: The accessible and recreational parks and protected areas (PAD-US-AR) dataset

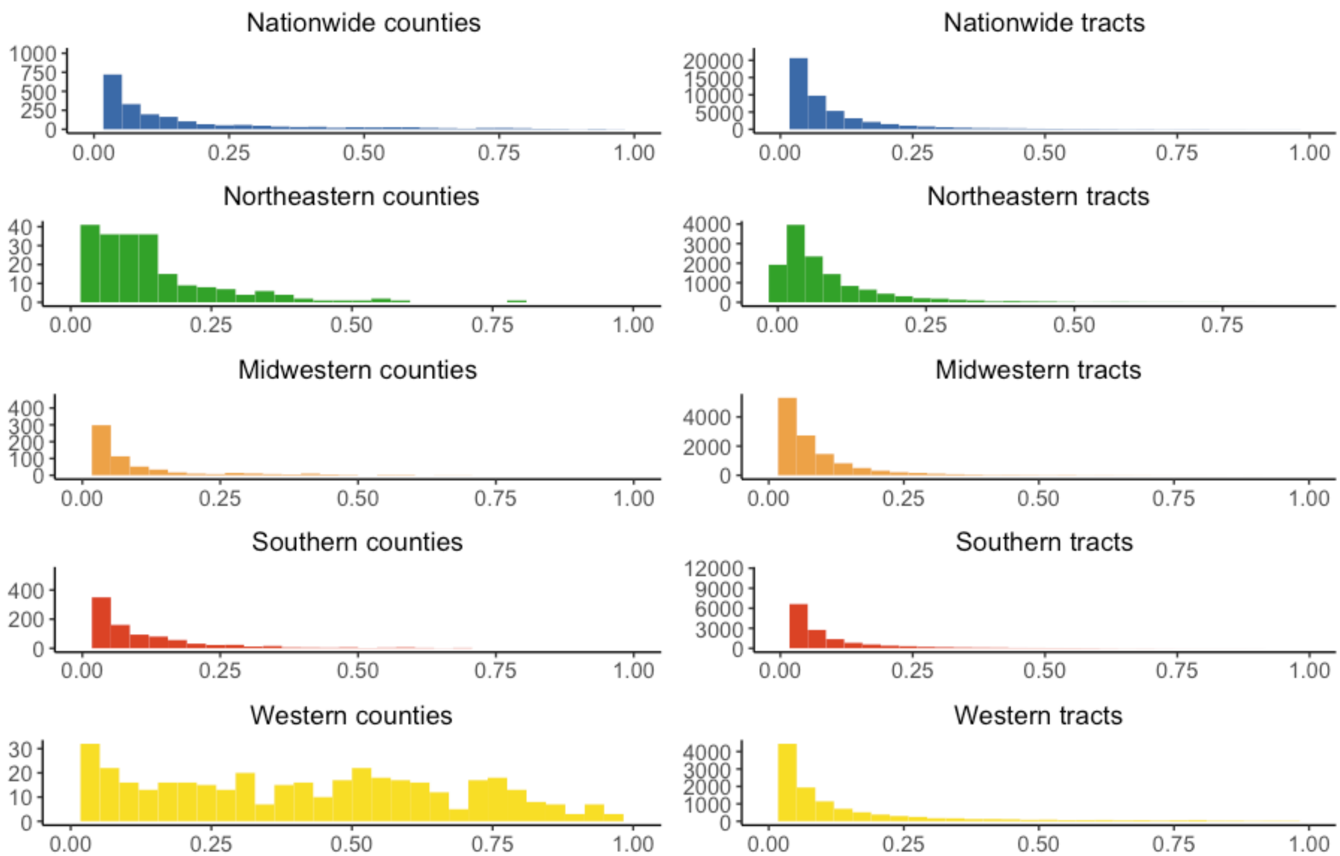
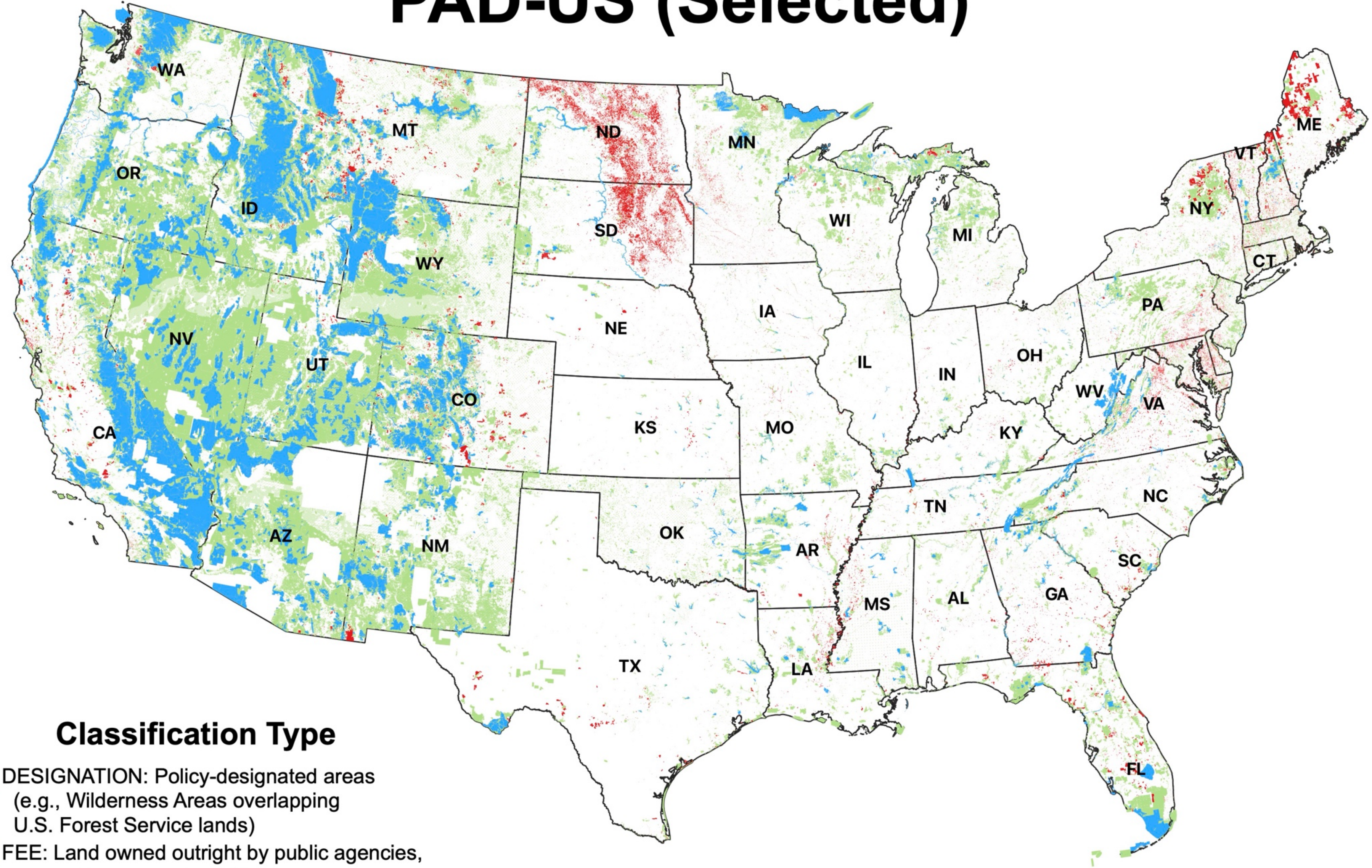


Figure S1. Histograms of county (left) and tract (right) level park cover estimates across the continental U.S.

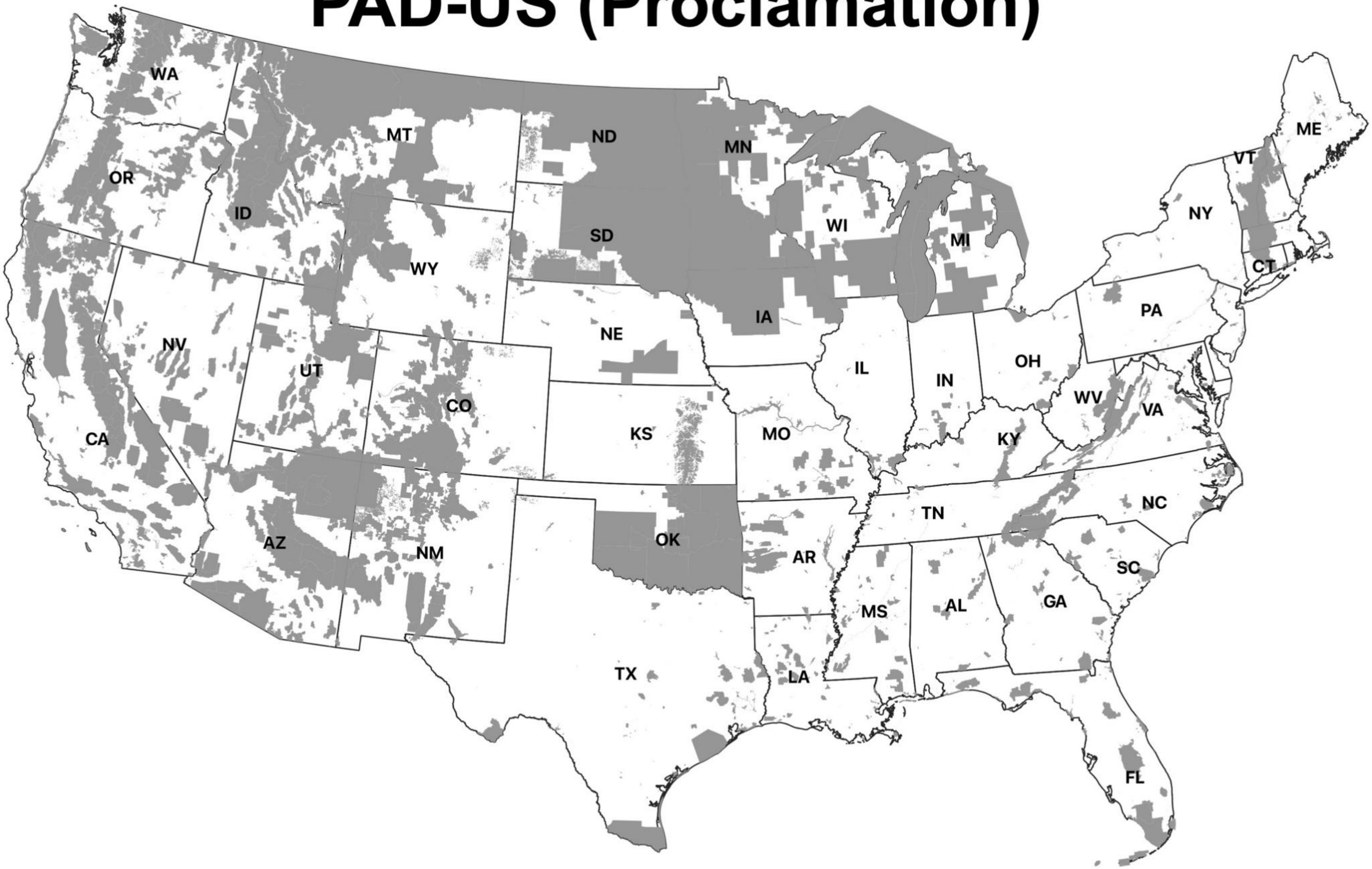
PAD-US (Selected)



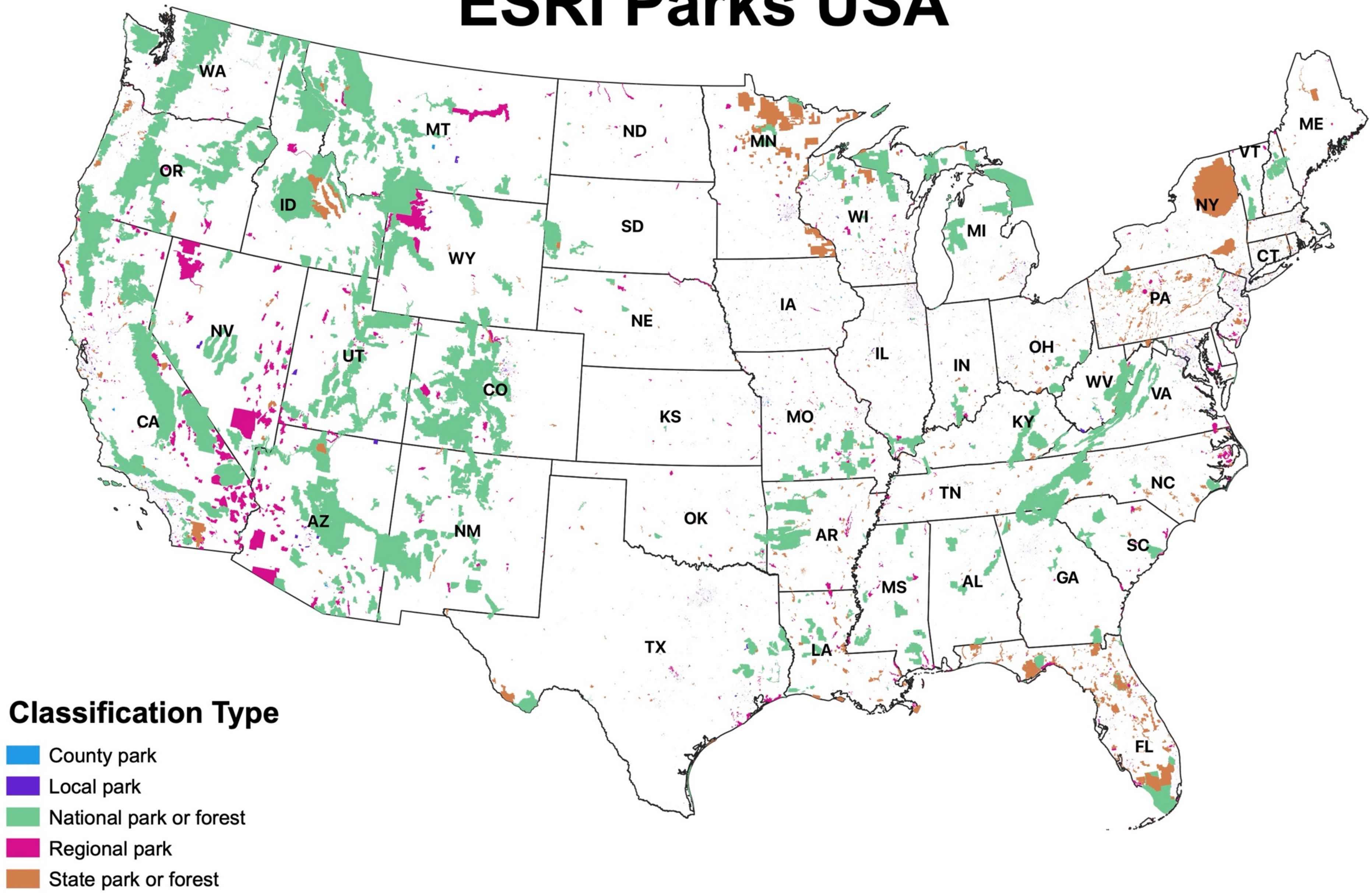
Classification Type

- DESIGNATION: Policy-designated areas (e.g., Wilderness Areas overlapping U.S. Forest Service lands)
- FEE: Land owned outright by public agencies, nonprofits, or private entities
- EASEMENT: Non-sensitive conservation and open space easements

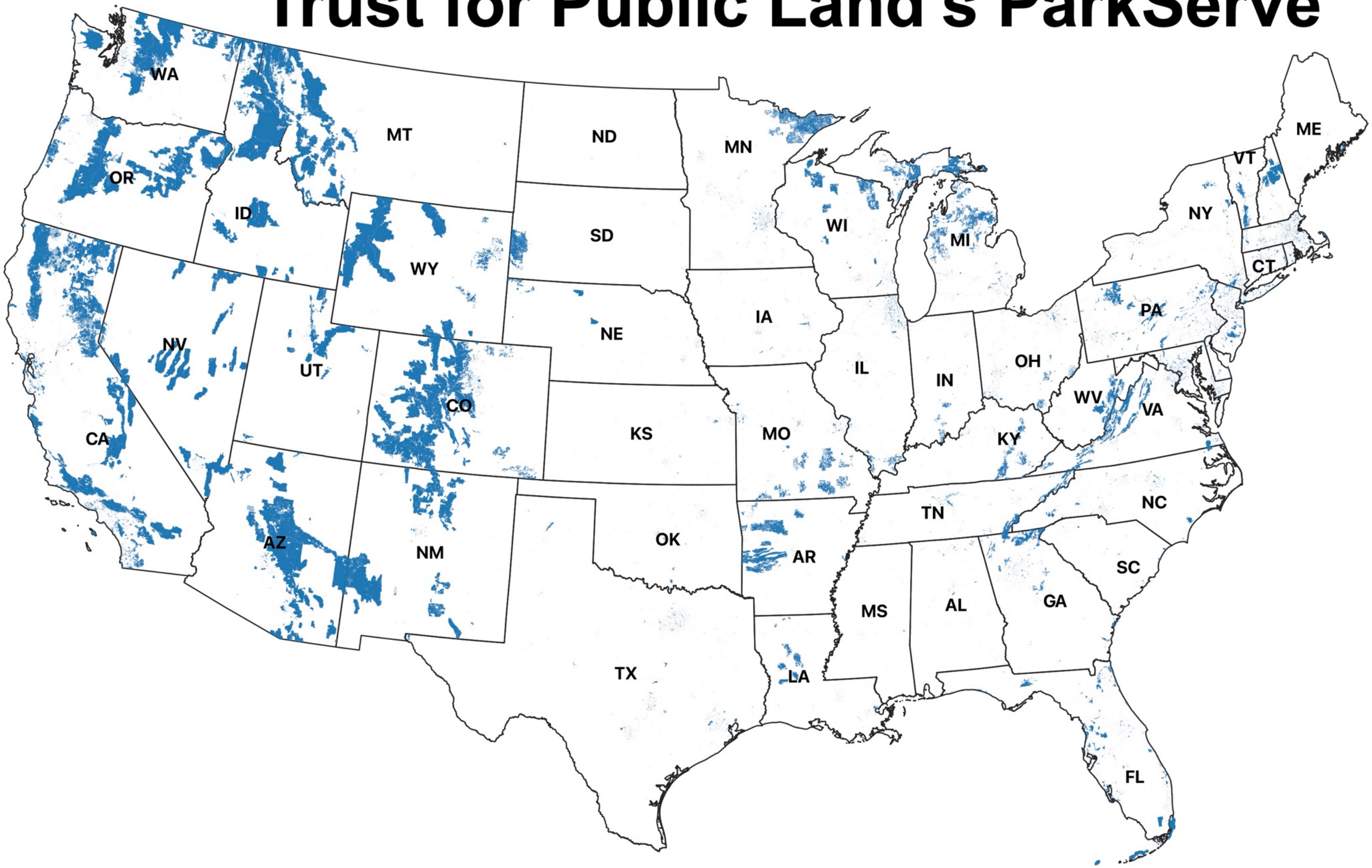
PAD-US (Proclamation)

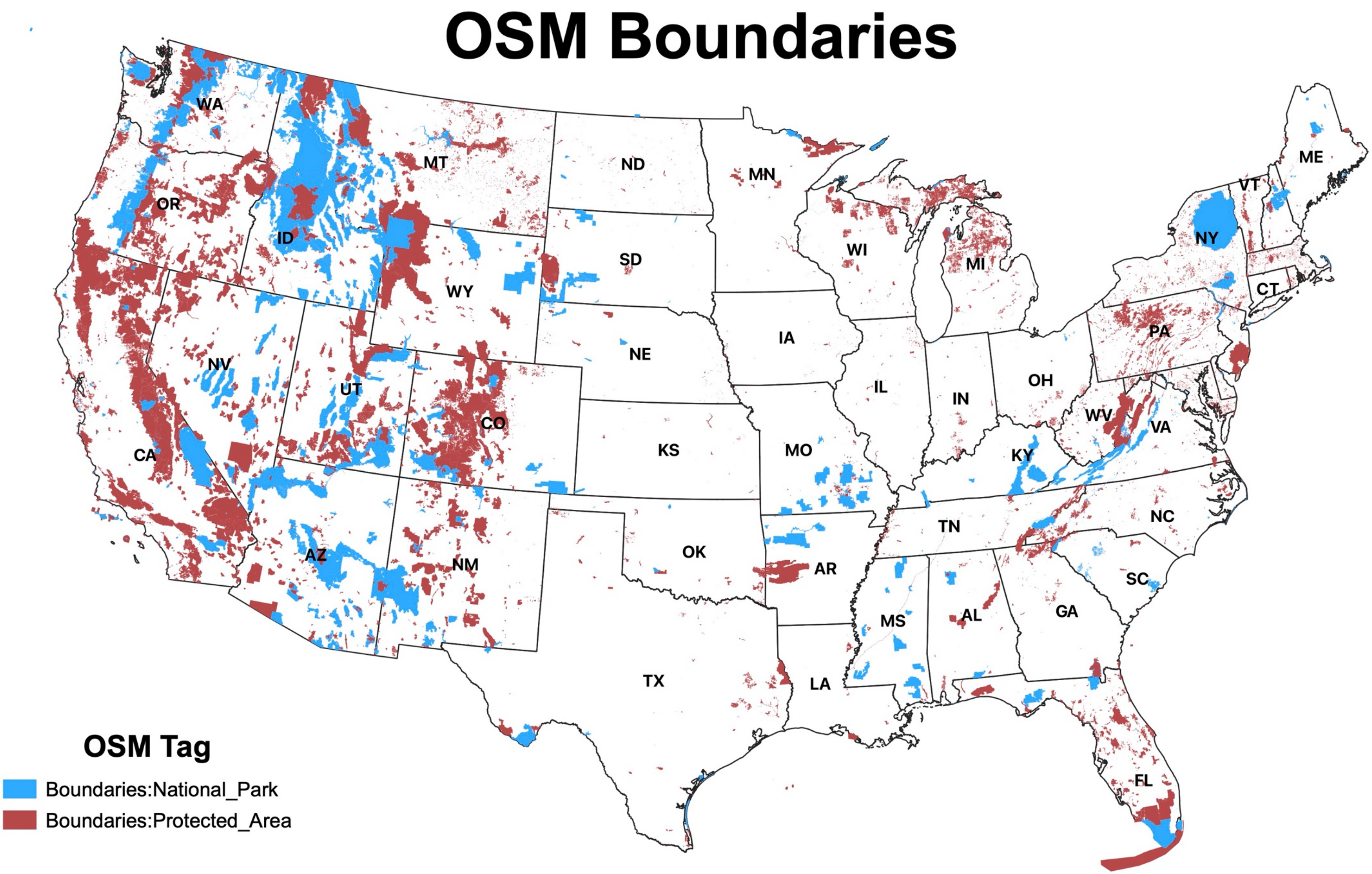


ESRI Parks USA



Trust for Public Land's ParkServe




OSM Boundaries
WA
OR
ID
MT
ND
MN
SD
WY
NV
UT
CO
NE
IA
WI
MI
IL
IN
OH
PA
NY
VT
ME
CT
WV
VA
KY
MO
KS
CA
AZ
NM
OK
AR
TN
NC
SC
GA
AL
MS
LA
TX
FL
OSM Tag
Boundaries:National_Park
Boundaries:Protected_Area

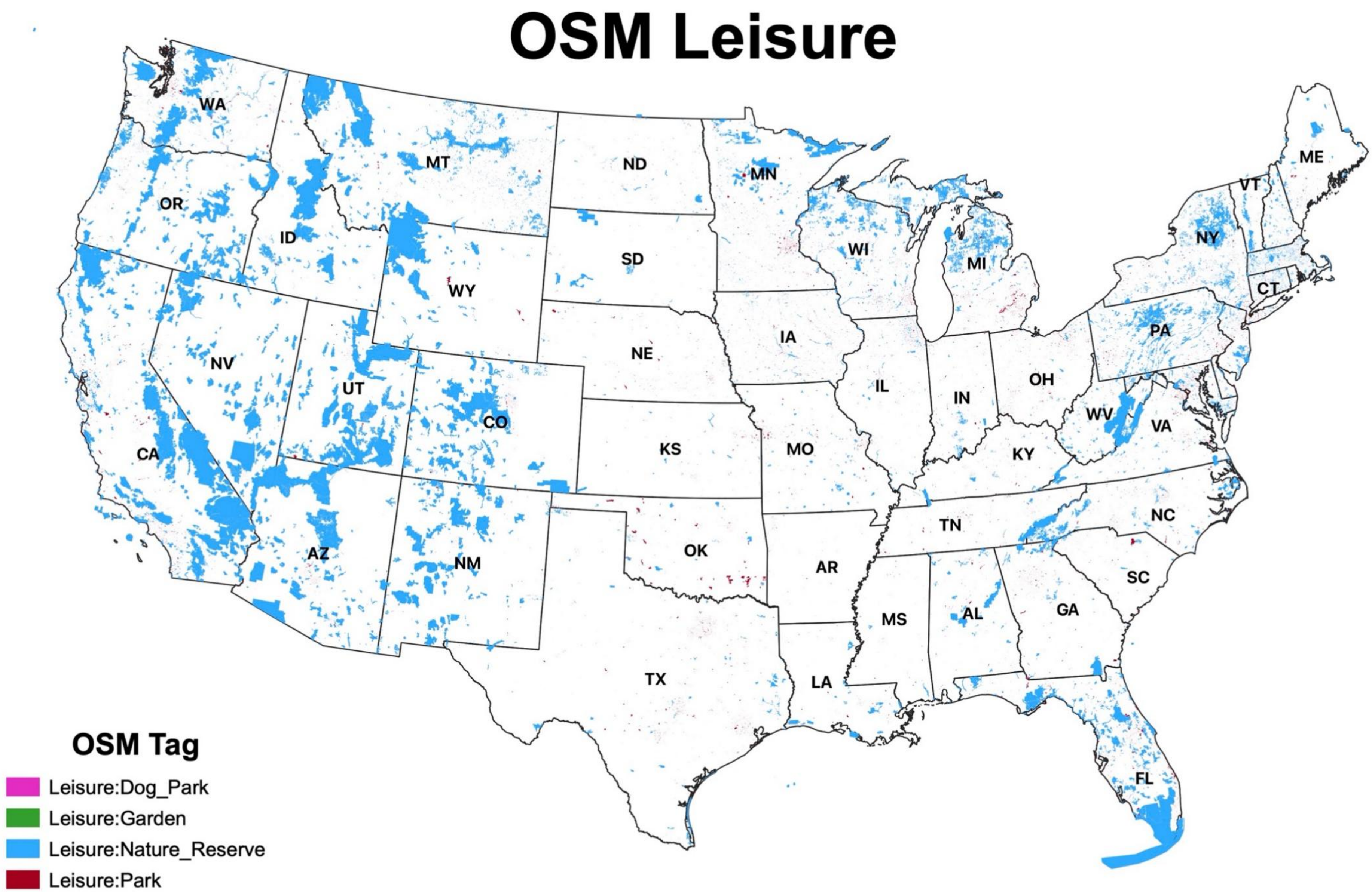


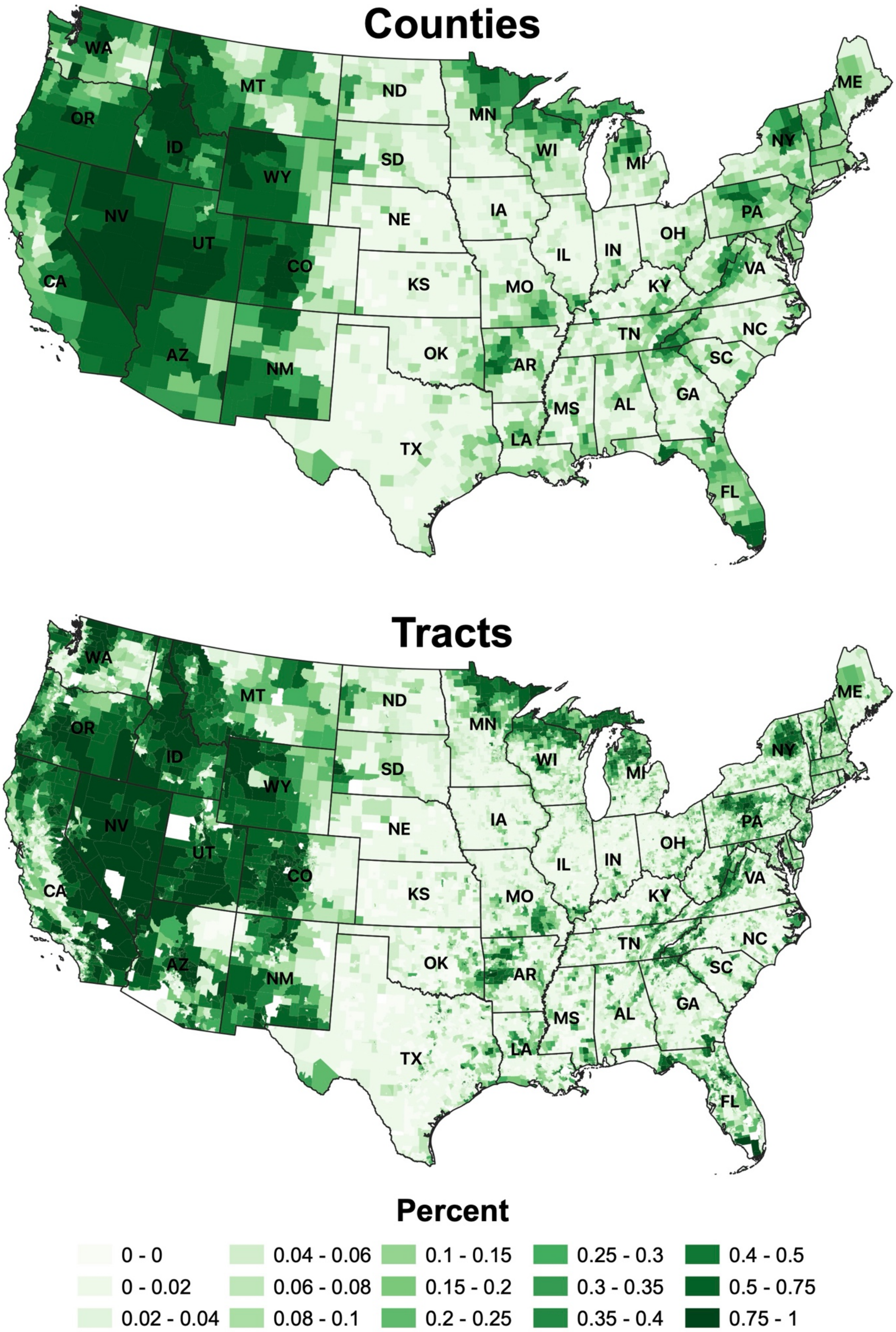
Figure S2. Preceding pages include nationwide comparisons between park cover datasets.

Table S1. Descriptive statistics for the PAD-US-AR and other green space measures

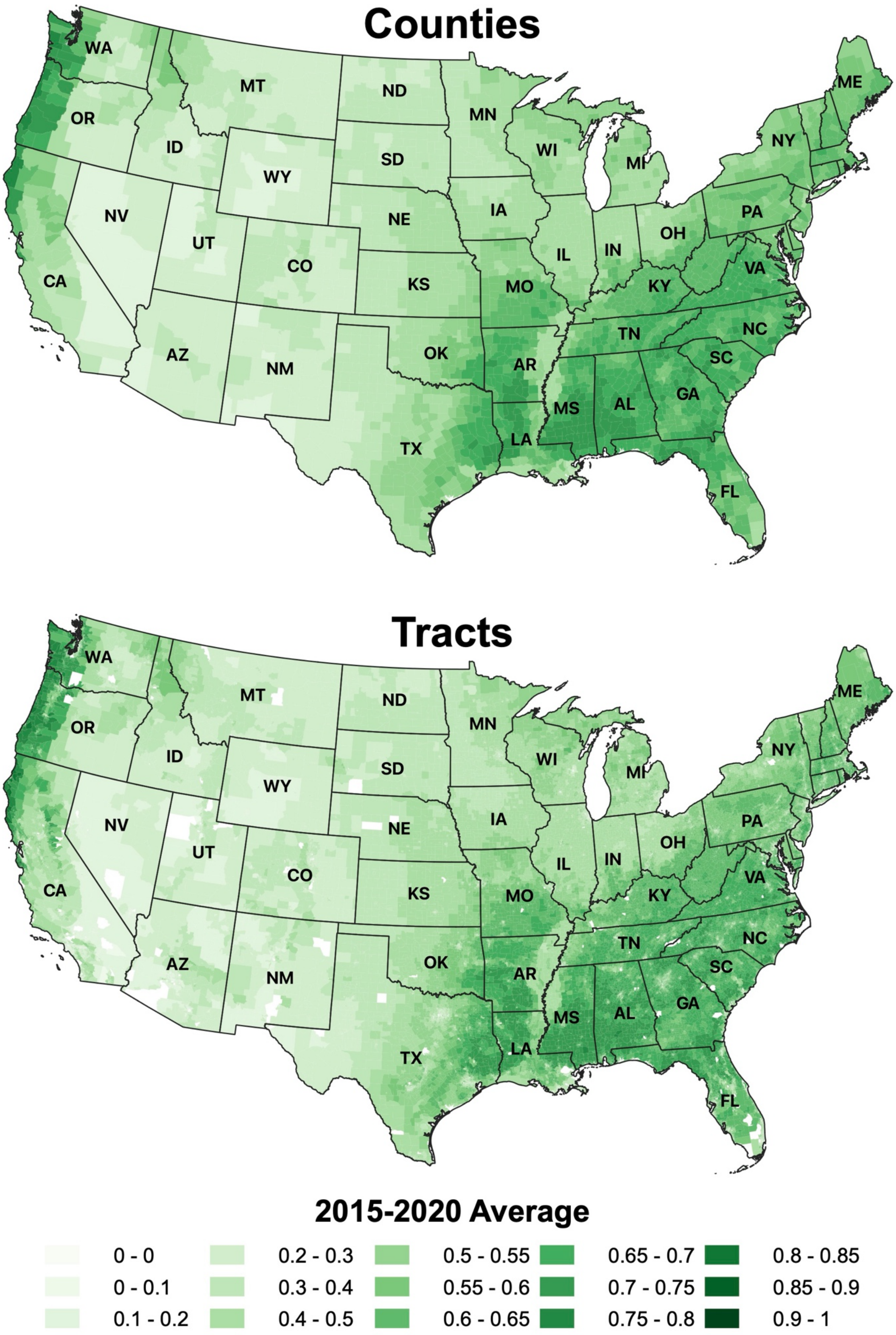
	Counties			Tracts		
Nationwide						
	Counties (N=3108)			**Tracts** (N=70378)		
	Med	*IQR*	*Range*	*Med*	IQR	*Range*
Public park cover	0.04	0.12	0.97 (0-0.97)	0.03	0.07	1 (0-1)
Tree canopy cover	0.27	0.47	0.87 (0-0.87)	0.13	0.3	0.93 (0-0.93)
NDVI annual average	0.53	0.22	0.66 (0.12-0.78)	0.46	0.21	0.81 (0.02-0.83)
NDVI summertime max	0.82	0.19	0.79 (0.13-0.93)	0.65	0.31	1.04 (-0.11-0.93)
Northeast						
	Counties (N=217)			**Tracts** (N=12882)		
	Med	*IQR*	*Range*	*Med*	IQR	*Range*
Public park cover	0.11	0.12	0.79 (0.01-0.79)	0.05	0.08	0.88 (0-0.88)
Tree canopy cover	0.53	0.24	0.75 (0.03-0.77)	0.24	0.37	0.83 (0-0.83)
NDVI annual average	0.58	0.05	0.43 (0.22-0.65)	0.49	0.23	0.64 (0.07-0.71)
NDVI summertime max	0.86	0.07	0.57 (0.35-0.93)	0.71	0.29	1.04 (-0.11-0.93)
Midwest						
	Counties (N=1055)			**Tracts** (N=16751)		
	Med	*IQR*	*Range*	*Med*	IQR	*Range*
Public park cover	0.02	0.05	0.7 (0-0.7)	0.04	0.07	0.88 (0-0.88)
Tree canopy cover	0.08	0.21	0.75 (0-0.75)	0.1	0.16	0.78 (0-0.78)
NDVI annual average	0.45	0.11	0.42 (0.25-0.66)	0.44	0.11	0.65 (0.02-0.67)
NDVI summertime max	0.84	0.1	0.72 (0.2-0.91)	0.71	0.21	0.79 (0.13-0.92)
South						
	Counties (N=1422)			**Tracts** (N=25579)		
	Med	*IQR*	*Range*	*Med*	IQR	*Range*
Public park cover	0.03	0.09	0.7 (0-0.7)	0.02	0.06	0.92 (0-0.92)
Tree canopy cover	0.47	0.37	0.87 (0-0.87)	0.26	0.36	0.93 (0-0.93)
NDVI annual average	0.63	0.11	0.57 (0.17-0.74)	0.55	0.17	0.69 (0.08-0.77)
NDVI summertime max	0.83	0.11	0.7 (0.21-0.91)	0.7	0.24	0.84 (0.09-0.93)
West						
	Counties (N=414)			**Tracts** (N=15166)		
	Med	*IQR*	*Range*	*Med*	IQR	*Range*
Public park cover	0.4	0.46	0.97 (0-0.97)	0.04	0.1	1 (0-1)
Tree canopy cover	0.09	0.21	0.66 (0-0.66)	0.02	0.06	0.74 (0-0.74)
NDVI annual average	0.31	0.15	0.66 (0.12-0.78)	0.31	0.17	0.78 (0.05-0.83)
NDVI summertime max	0.46	0.24	0.74 (0.13-0.88)	0.36	0.22	0.85 (0.05-0.91)

Note: Med.=median, IQR=interquartile range

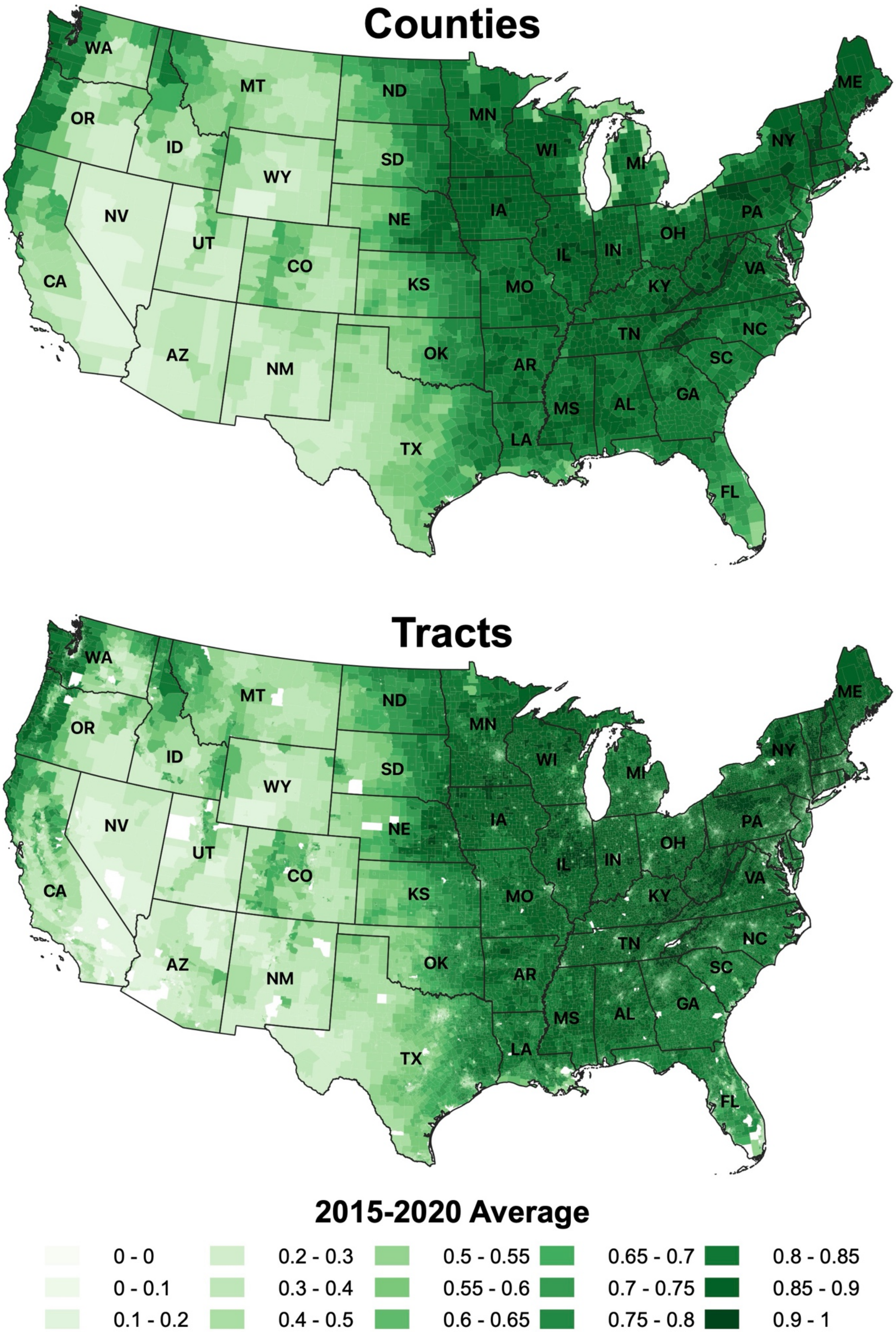
Public Park Cover (PAD-US-AR)



Annual NDVI


Counties
WA
MT
ND
MN
ME
OR
ID
SD
WI
NY
MI
WY
NV
UT
NE
IA
PA
OH
IL
IN
CO
CA
KS
MO
VA
KY
TN
NC
OK
AZ
NM
AR
SC
GA
MS
AL
LA
TX
FL
Tracts
WA
MT
ND
MN
ME
OR
ID
SD
WI
NY
MI
WY
NV
UT
NE
IA
PA
OH
IL
IN
CO
CA
KS
MO
VA
KY
TN
NC
OK
AZ
NM
AR
SC
GA
MS
AL
LA
TX
FL
2015-2020 Average
0 - 0
0 - 0.1
0.1 - 0.2
0.2 - 0.3
0.3 - 0.4
0.4 - 0.5
0.5 - 0.55
0.55 - 0.6
0.6 - 0.65
0.65 - 0.7
0.7 - 0.75
0.75 - 0.8
0.8 - 0.85
0.85 - 0.9
0.9 - 1

Summertime Maximum NDVI



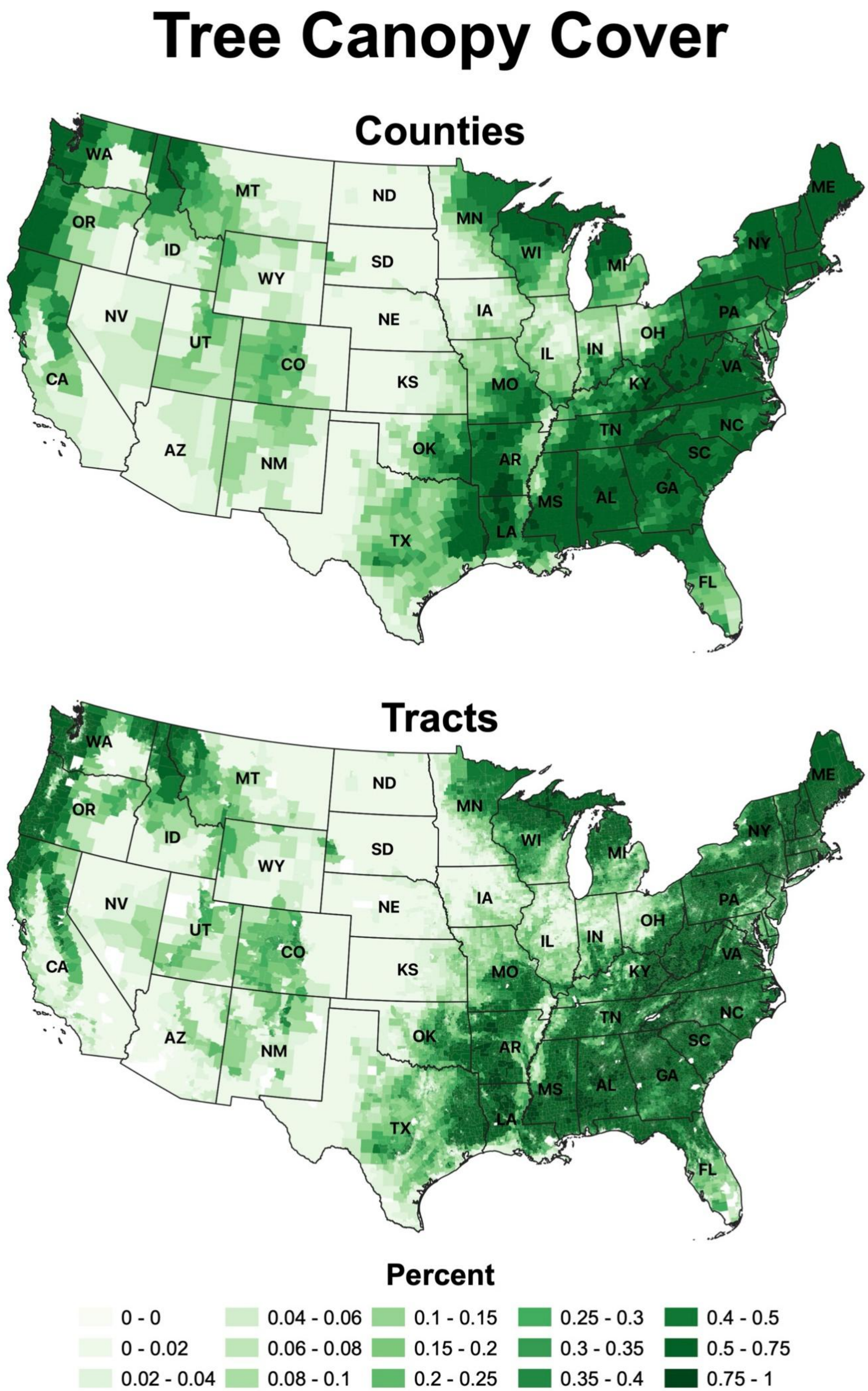


Figure S3. Maps of public park cover and other green space measures in counties and tracts across the U.S.

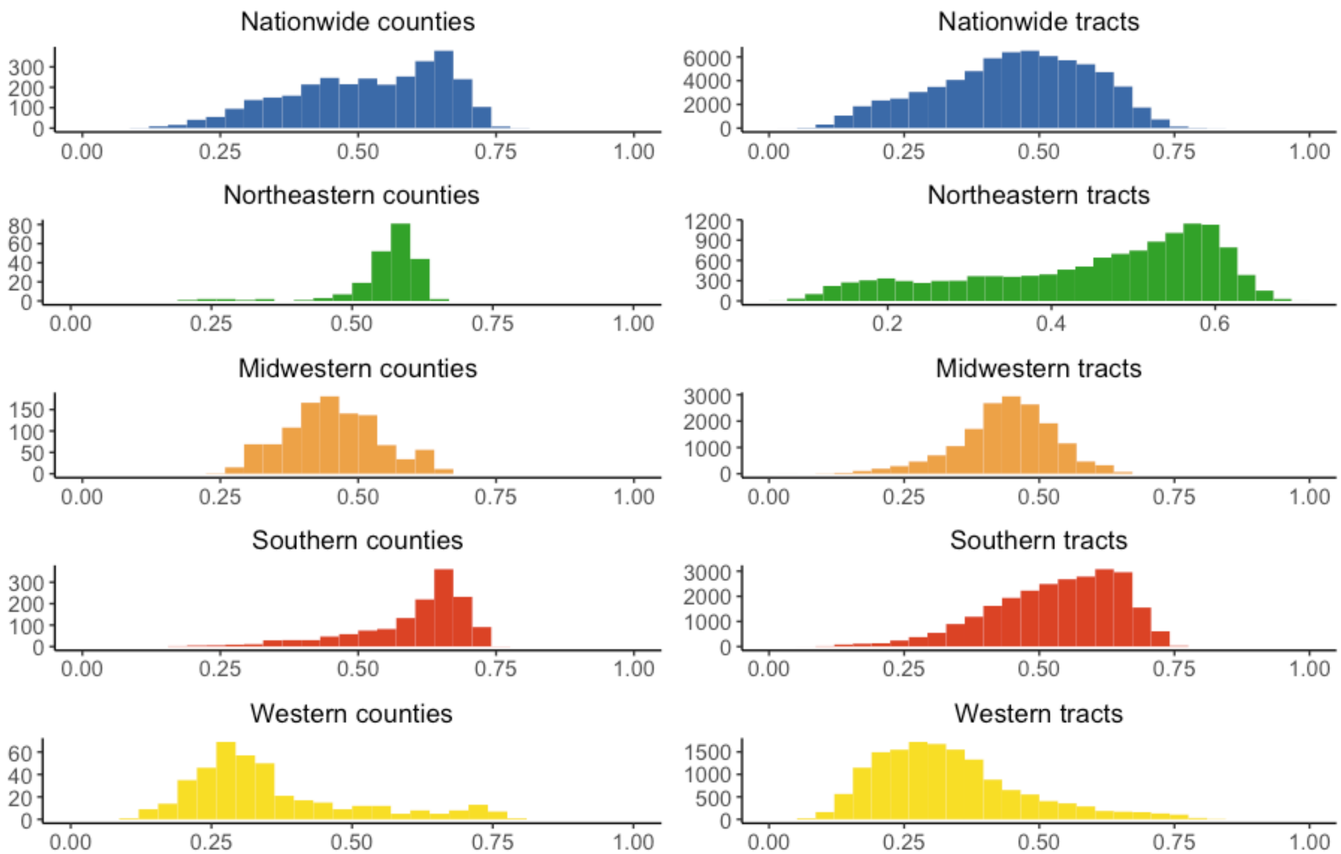


Figure S4. Histograms of county-level and tract-level **NDVI annual averages** across the continental U.S. and within census regions

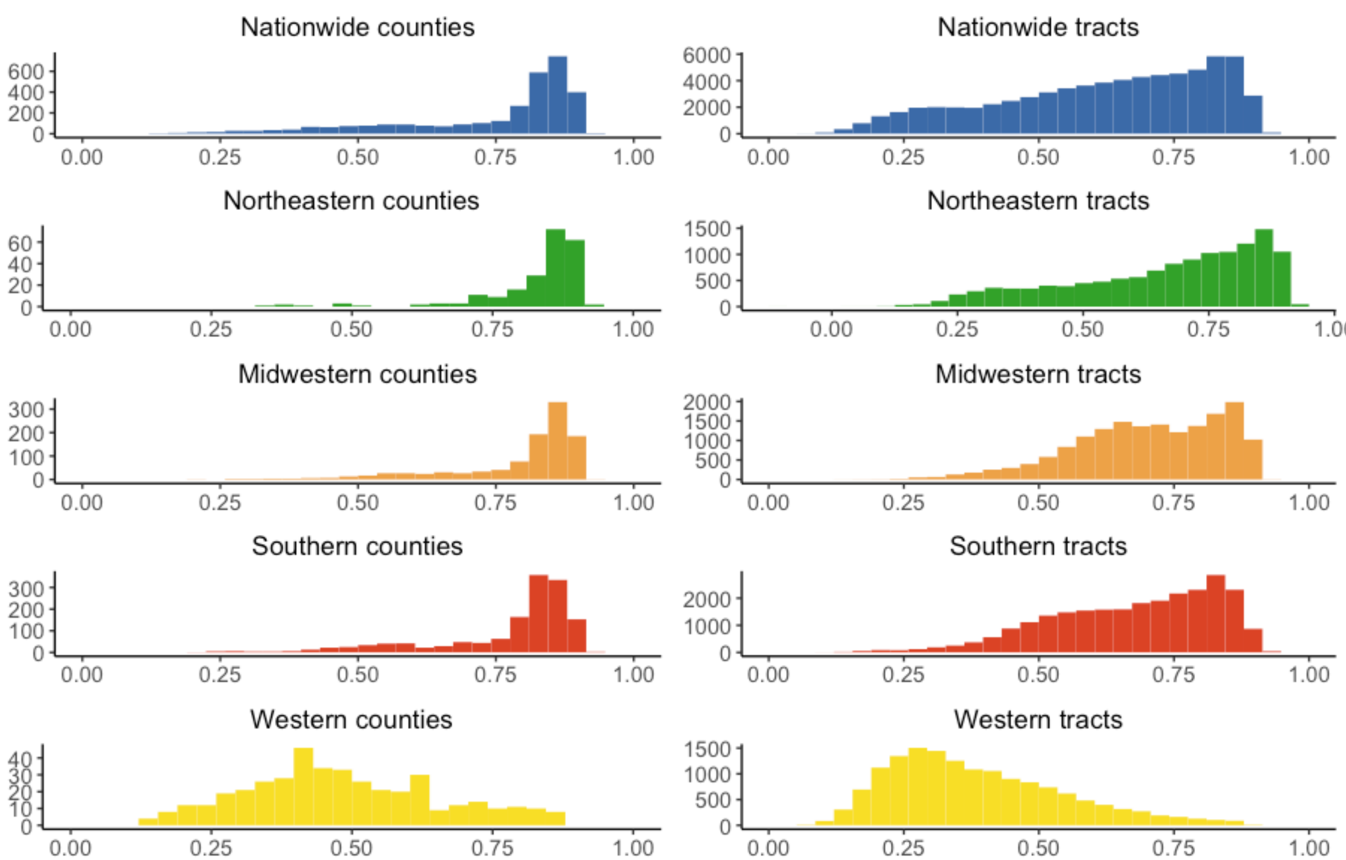


Figure S5. Histograms of county-level and tract-level **NDVI summertime maximums** across the continental U.S. and within census regions

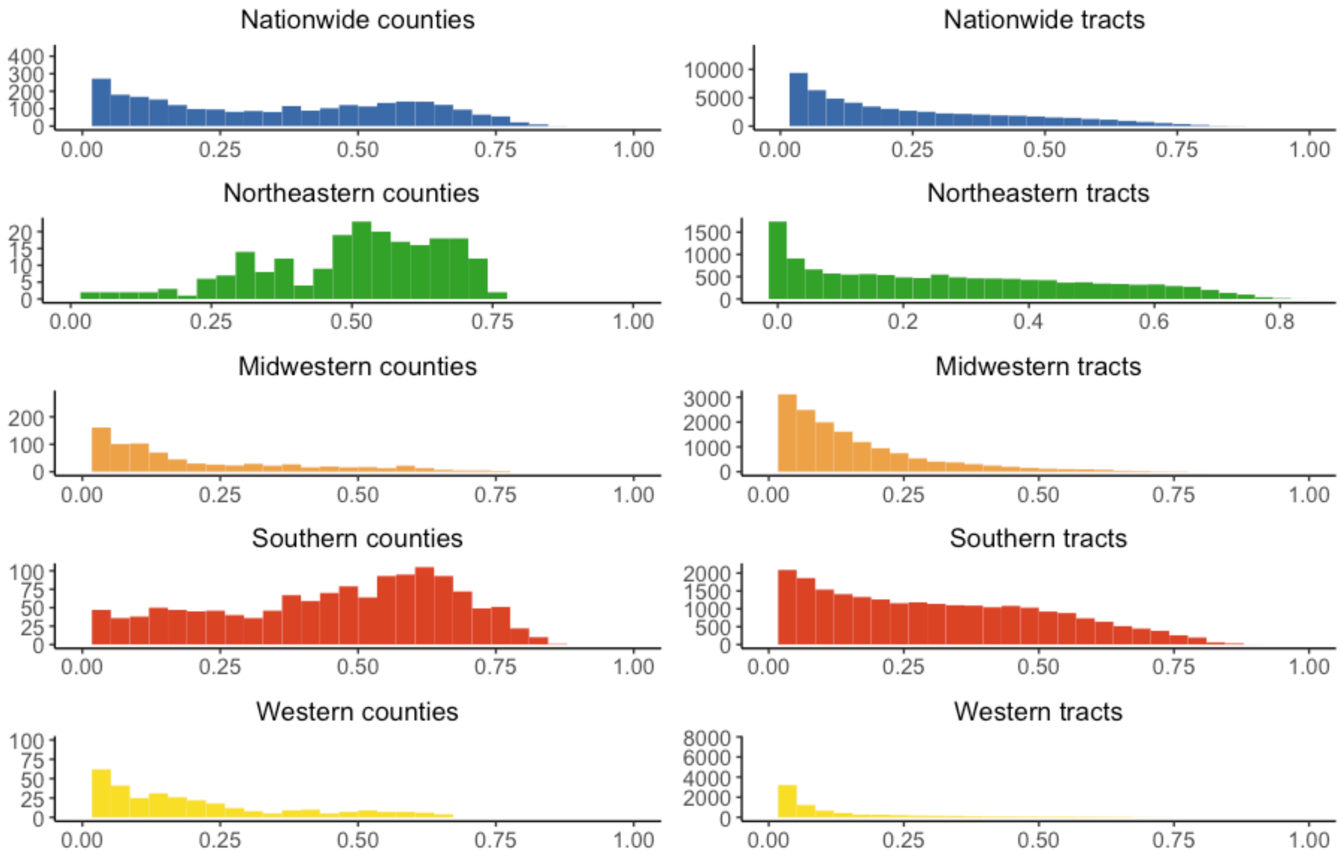


Figure S6. Histograms of county-level and tract-level **tree canopy cover** across the continental U.S. and within census regions

Table S2. Sociodemographic characteristics examined in this study.

Variable Name	Description	Source
Population density	Number of people per km^2	American Community Survey (ACS) 5-Year Data: 2015-2019
Median home value	Median value (USD) of owner-occupied housing units	ACS 2015-2019
% poverty	Ratio of households with incomes below the poverty level in the past 12 months to total number of households	ACS 2015-2019
GINI index	Gini index of income inequality, ranging from 0 (perfect equality, where everyone receives an equal share) to 1 (perfect inequality, where only one recipient or group receives all the income)	ACS 2015-2019
% high school degree	Ratio of high school diploma earners (or greater levels of educational achievement) among people 25 years or over to total population 25 years or over	ACS 2015-2019
% college degree	Ratio of Bachelor's degree holders (or greater levels of educational achievement) among people 25 years or over to total population 25 years or over	ACS 2015-2019
% unemployed	Ratio of people aged 16 years or over in the labor force to total population aged 16 years or over	ACS 2015-2019
% employed natural resources	Ratio of people aged 16 years or over in agriculture, forestry, fishing and hunting, and mining industries to total population aged 16 years or over	ACS 2015-2019
% NH Black	Ratio of population who are not Hispanic or Latino and Black or African American to total population	ACS 2015-2019
% NH Asian	Ratio of population who are not Asian to total population	ACS 2015-2019
% Hispanic	Ratio of Hispanic or Latino Origin by Race to Total Population	ACS 2015-2019
% 65+ years	Ratio of population who are 65 years or over to total population	ACS 2015-2019
% Female	Ratio of population who are female to total population	ACS 2015-2019
Total population	Population size (number of residents)	ACS 2015-2019

Table S3. Descriptive statistics for socio-demographic factors (**Nationwide**)

	Counties (N=3107)			Tracts (N=70378)		
	Med	*IQR*	*Range*	*Med*	IQR	*Range*
Population density	17.01	38.3	27648.58 (0.06-27648.64)	874.93	1941.9	99783.32 (0.05-99783.37)
Median household income	51658	15460	120795 (21504-142299)	59725.5	37168.25	245872 (4129-250001)
Median home value	127200	76300	1073400 (24400-1097800)	191800	213175	1990002 (9999-2000001)
% poverty	0.14	0.07	0.46 (0.02-0.48)	0.11	0.12	0.8 (0-0.8)
Gini index	0.44	0.05	0.39 (0.32-0.71)	0.42	0.08	0.77 (0.05-0.82)
% high school degree	0.88	0.08	0.72 (0.26-0.99)	0.9	0.12	0.76 (0.24-1)
% college degree	0.2	0.11	0.78 (0-0.78)	0.26	0.26	0.97 (0-0.97)
% unemployed	0.41	0.11	0.65 (0.2-0.85)	0.36	0.12	0.96 (0.03-0.99)
% employed natural resources	0.04	0.07	0.6 (0-0.6)	0.01	0.02	0.67 (0-0.67)
% NH Black	0.02	0.09	0.87 (0-0.87)	0.04	0.14	1 (0-1)
% NH Asian	0.01	0.01	0.36 (0-0.36)	0.02	0.05	0.94 (0-0.94)
% Hispanic	0.04	0.08	0.99 (0-0.99)	0.08	0.18	1 (0-1)
% female	0.5	0.02	0.3 (0.27-0.57)	0.51	0.04	0.94 (0.01-0.95)
% 65+ years	0.18	0.05	0.54 (0.03-0.57)	0.16	0.09	0.92 (0-0.92)
Total population	25946	56714	10081472 (98-10081570)	4167	2615	71534 (507-72041)

Table S4. Descriptive statistics for socio-demographic factors (**Northeast**)

	Counties (N=217)			Tracts (N=12882)		
	Med	*IQR*	*Range*	*Med*	IQR	*Range*
Population density	63.09	163.18	27647.68 (0.96-27648.64)	1245.24	3862.43	84143.01 (0.07-84143.08)
Median home value	177100	118100	1008300 (76400-1084700)	265300	275375	1990002 (9999-2000001)
% poverty	0.12	0.04	0.24 (0.04-0.28)	0.09	0.11	0.72 (0-0.72)
Gini index	0.44	0.03	0.2 (0.39-0.6)	0.42	0.08	0.52 (0.25-0.77)
% high school degree	0.91	0.04	0.23 (0.73-0.96)	0.92	0.1	0.68 (0.32-1)
% college degree	0.28	0.15	0.53 (0.08-0.61)	0.32	0.27	0.96 (0-0.97)
% unemployed	0.38	0.07	0.58 (0.27-0.85)	0.35	0.1	0.87 (0.06-0.93)
% employed natural resources	0.02	0.02	0.12 (0-0.12)	0	0.01	0.31 (0-0.31)
% NH Black	0.03	0.05	0.41 (0-0.41)	0.03	0.11	0.99 (0-0.99)
% NH Asian	0.01	0.03	0.25 (0-0.25)	0.02	0.06	0.92 (0-0.92)
% Hispanic	0.03	0.06	0.55 (0.01-0.56)	0.06	0.14	0.97 (0-0.97)
% female	0.51	0.01	0.26 (0.27-0.53)	0.51	0.04	0.79 (0.16-0.95)
% 65+ years	0.19	0.04	0.19 (0.12-0.3)	0.17	0.08	0.92 (0-0.92)
Total population	102642	245889	2585459 (4515-2589974)	3973	2409.5	27593 (516-28109)

Table S5. Descriptive statistics for socio-demographic factors (**Midwest**)

	Counties (N=1055)			**Tracts** (N=16751)		
	Med	*IQR*	*Range*	*Med*	IQR	*Range*
Population density	12.29	27.62	2095.15 (0.18-2095.33)	653.58	1502.64	99783.24 (0.14-99783.37)
Median home value	118400	54600	287200 (26000-313200)	145800	107850	1538501 (9999-1548500)
% poverty	0.12	0.05	0.46 (0.02-0.48)	0.11	0.11	0.8 (0-0.8)
Gini index	0.43	0.04	0.22 (0.33-0.56)	0.41	0.07	0.71 (0.11-0.82)
% high school degree	0.91	0.04	0.41 (0.57-0.98)	0.92	0.08	0.72 (0.28-1)
% college degree	0.2	0.08	0.51 (0.09-0.59)	0.23	0.22	0.96 (0-0.96)
% unemployed	0.37	0.08	0.4 (0.24-0.64)	0.35	0.11	0.92 (0.08-0.99)
% employed natural resources	0.05	0.08	0.51 (0-0.51)	0.01	0.02	0.55 (0-0.55)
% NH Black	0.01	0.02	0.46 (0-0.46)	0.03	0.1	1 (0-1)
% NH Asian	0.01	0.01	0.14 (0-0.14)	0.01	0.03	0.86 (0-0.86)
% Hispanic	0.03	0.03	0.61 (0-0.61)	0.04	0.06	0.99 (0-0.99)
% female	0.5	0.01	0.19 (0.37-0.55)	0.51	0.04	0.67 (0.03-0.7)
% 65+ years	0.19	0.05	0.29 (0.07-0.36)	0.16	0.08	0.57 (0-0.57)
Total population	19941	36794.5	5197880 (395-5198275)	3718	2316.5	45968 (514-46482)

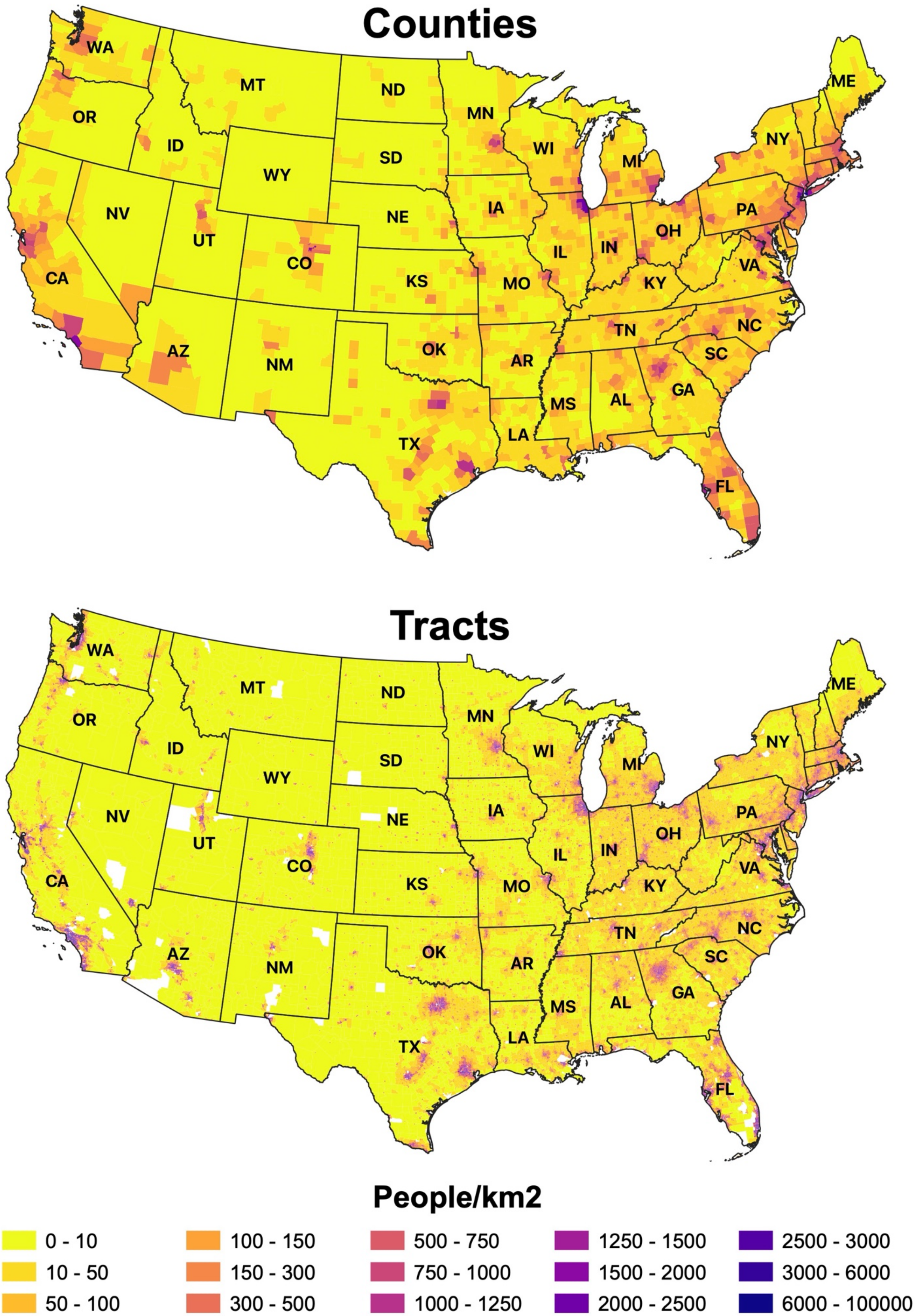
Table S6. Descriptive statistics for socio-demographic factors (**South**)

	Counties (N=1422)			**Tracts** (N=25579)		
	Med	*IQR*	*Range*	*Med*	IQR	*Range*
Population density	21.1	42.13	4300.39 (0.06-4300.45)	575.73	1353.27	30886.44 (0.15-30886.58)
Median home value	116400	70300	764900 (24400-789300)	156700	139150	1990002 (9999-2000001)
% poverty	0.17	0.08	0.4 (0.03-0.43)	0.13	0.13	0.79 (0-0.79)
Gini index	0.46	0.04	0.39 (0.32-0.71)	0.43	0.08	0.73 (0.05-0.79)
% high school degree	0.84	0.08	0.72 (0.26-0.99)	0.88	0.13	0.75 (0.25-1)
% college degree	0.17	0.1	0.78 (0-0.78)	0.23	0.25	0.95 (0-0.95)
% unemployed	0.45	0.1	0.58 (0.2-0.78)	0.38	0.14	0.96 (0.03-0.99)
% employed natural resources	0.03	0.05	0.51 (0-0.51)	0.01	0.03	0.51 (0-0.51)
% NH Black	0.09	0.24	0.87 (0-0.87)	0.1	0.25	1 (0-1)
% NH Asian	0.01	0.01	0.2 (0-0.2)	0.01	0.03	0.76 (0-0.76)
% Hispanic	0.05	0.09	0.99 (0-0.99)	0.07	0.17	1 (0-1)
% female	0.51	0.02	0.24 (0.33-0.57)	0.51	0.04	0.79 (0.01-0.8)
% 65+ years	0.18	0.05	0.54 (0.03-0.57)	0.16	0.09	0.89 (0-0.89)
Total population	26241.5	50747.25	4646532 (98-4646630)	4338	2898	71534 (507-72041)

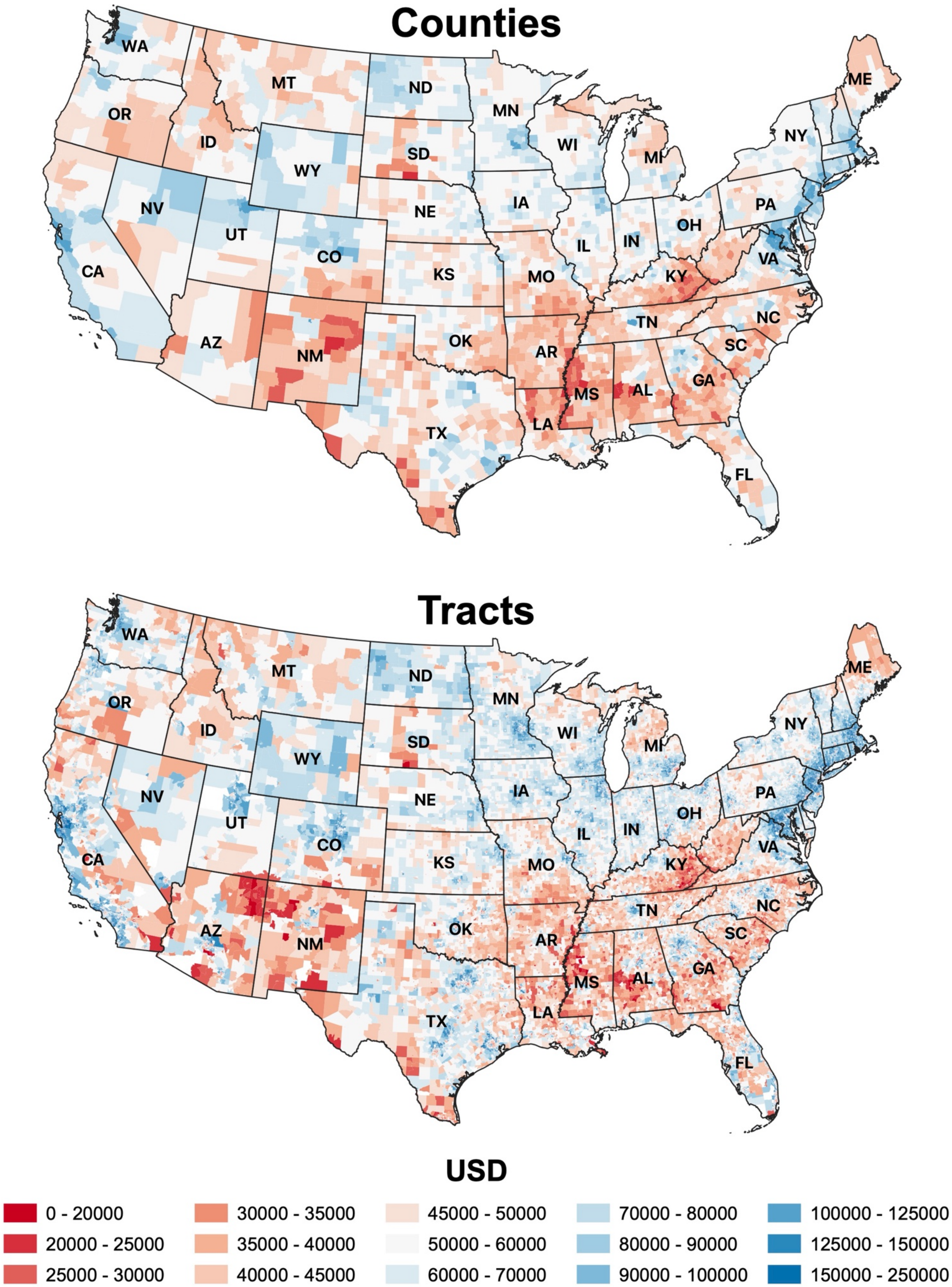
Table S7. Descriptive statistics for socio-demographic factors (**West**)

	Counties (N=414)			**Tracts** (N=15166)		
	Med	*IQR*	*Range*	*Med*	IQR	*Range*
Population density	4.3	17.74	7129.98 (0.08-7130.06)	1681.72	2659.41	50191.15 (0.05-50191.2)
Median home value	205000	129525	1037900 (59900-1097800)	355800	328350	1990002 (9999-2000001)
% poverty	0.13	0.06	0.31 (0.03-0.35)	0.1	0.11	0.8 (0-0.8)
Gini index	0.44	0.05	0.29 (0.32-0.61)	0.41	0.08	0.48 (0.23-0.72)
% high school degree	0.9	0.06	0.39 (0.6-0.99)	0.9	0.15	0.76 (0.24-1)
% college degree	0.23	0.13	0.6 (0.08-0.67)	0.28	0.29	0.96 (0-0.96)
% unemployed	0.41	0.11	0.49 (0.2-0.69)	0.36	0.11	0.92 (0.04-0.96)
% employed natural resources	0.07	0.12	0.59 (0-0.6)	0.01	0.02	0.67 (0-0.67)
% NH Black	0.01	0.01	0.14 (0-0.14)	0.02	0.05	0.85 (0-0.85)
% NH Asian	0.01	0.01	0.36 (0-0.36)	0.04	0.1	0.94 (0-0.94)
% Hispanic	0.12	0.21	0.84 (0-0.84)	0.21	0.35	1 (0-1)
% female	0.5	0.02	0.22 (0.34-0.56)	0.5	0.04	0.63 (0.05-0.67)
% 65+ years	0.19	0.09	0.32 (0.07-0.39)	0.14	0.09	0.91 (0-0.91)
Total population	23648.5	82103.75	10081129 (441-10081570)	4569.5	2536	30654 (507-31161)

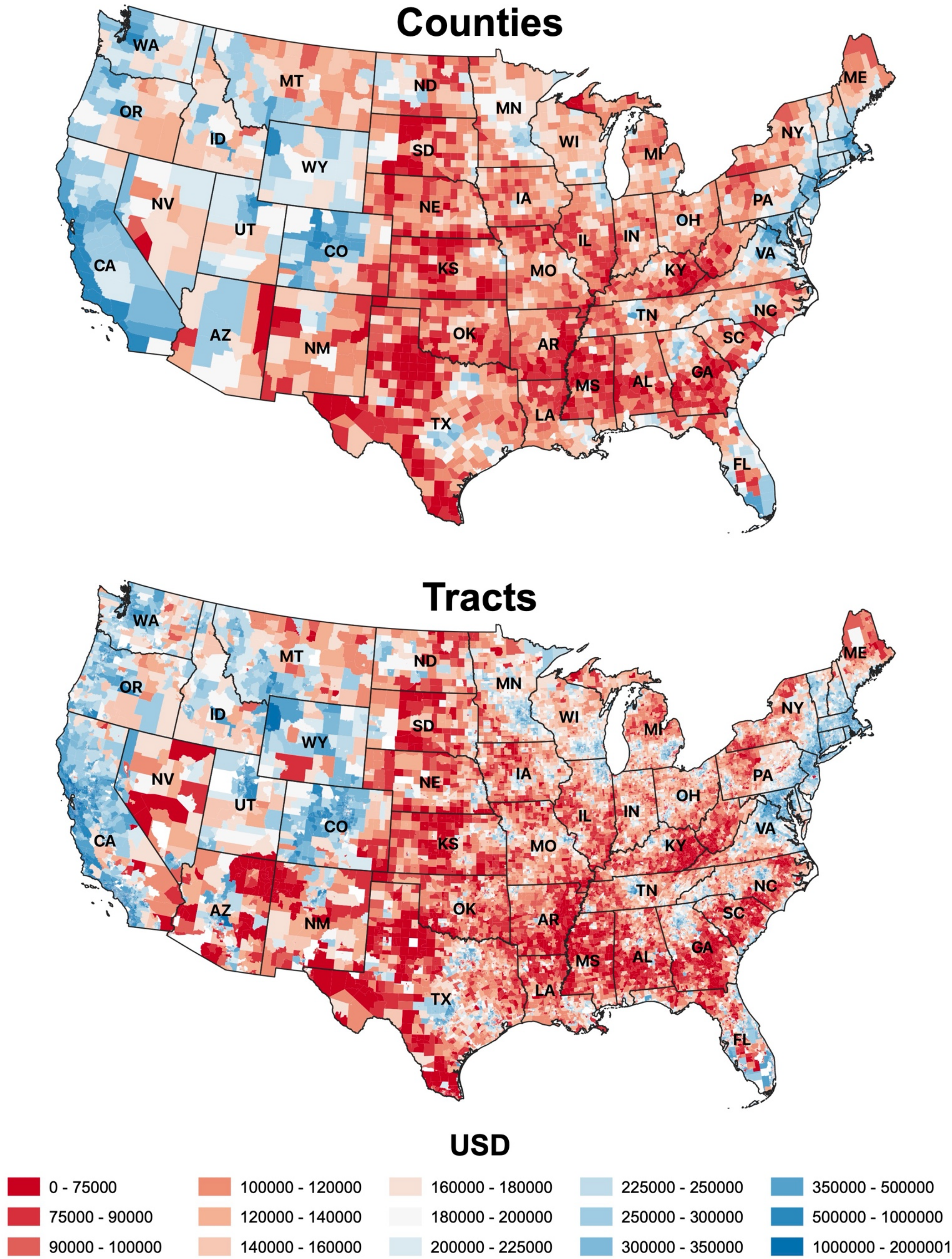
Population density



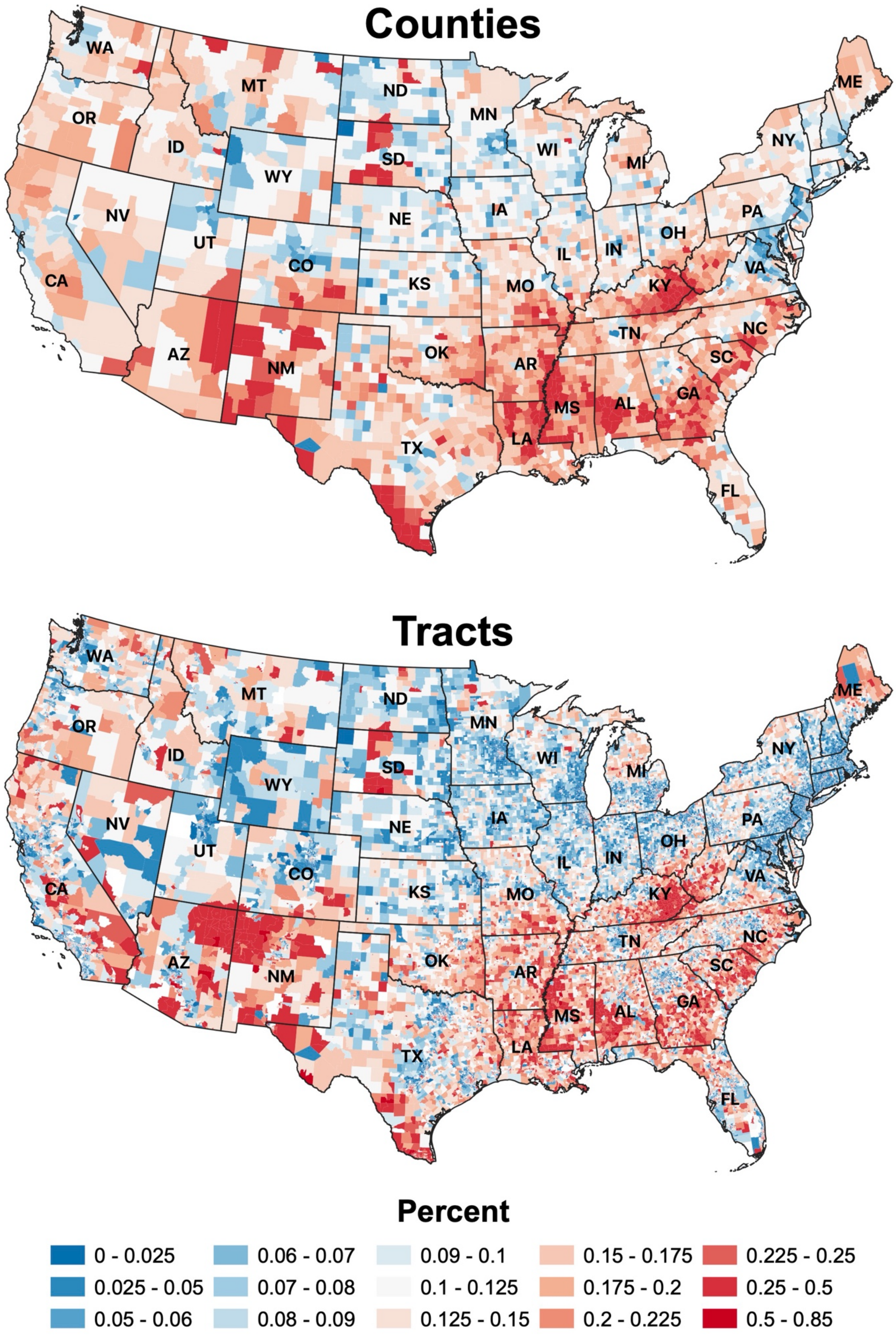
Median Household Income



Median Home Value

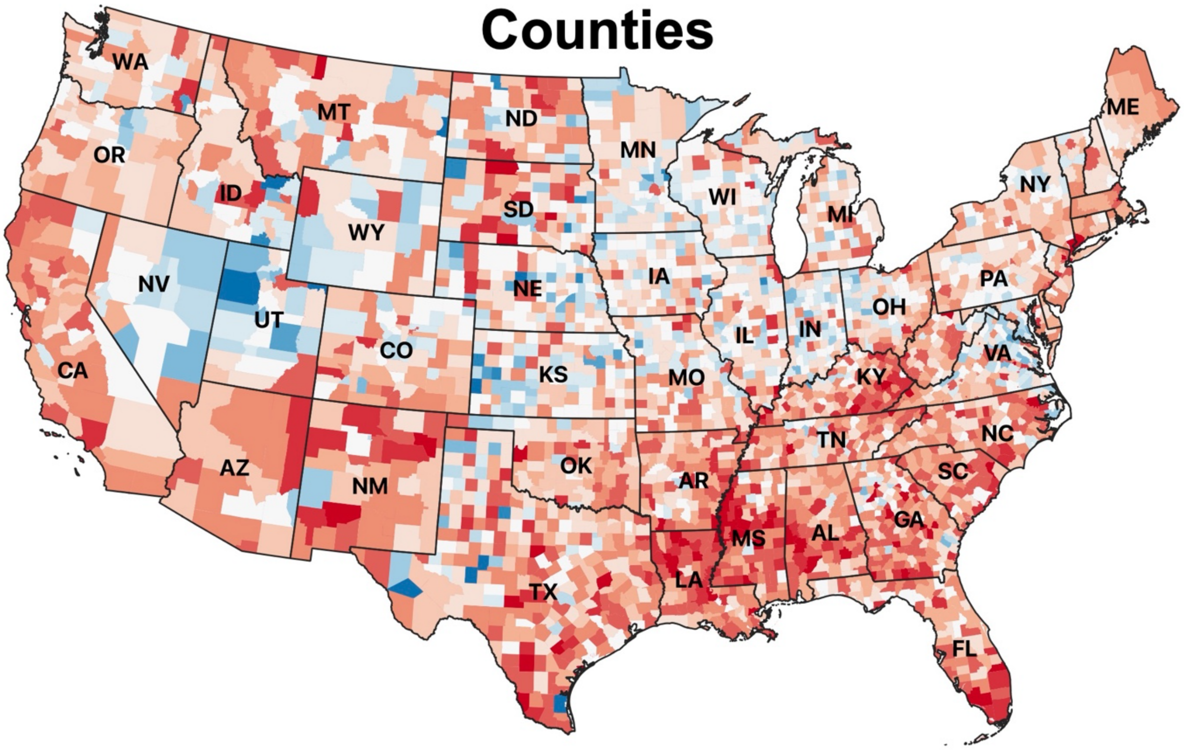


Poverty

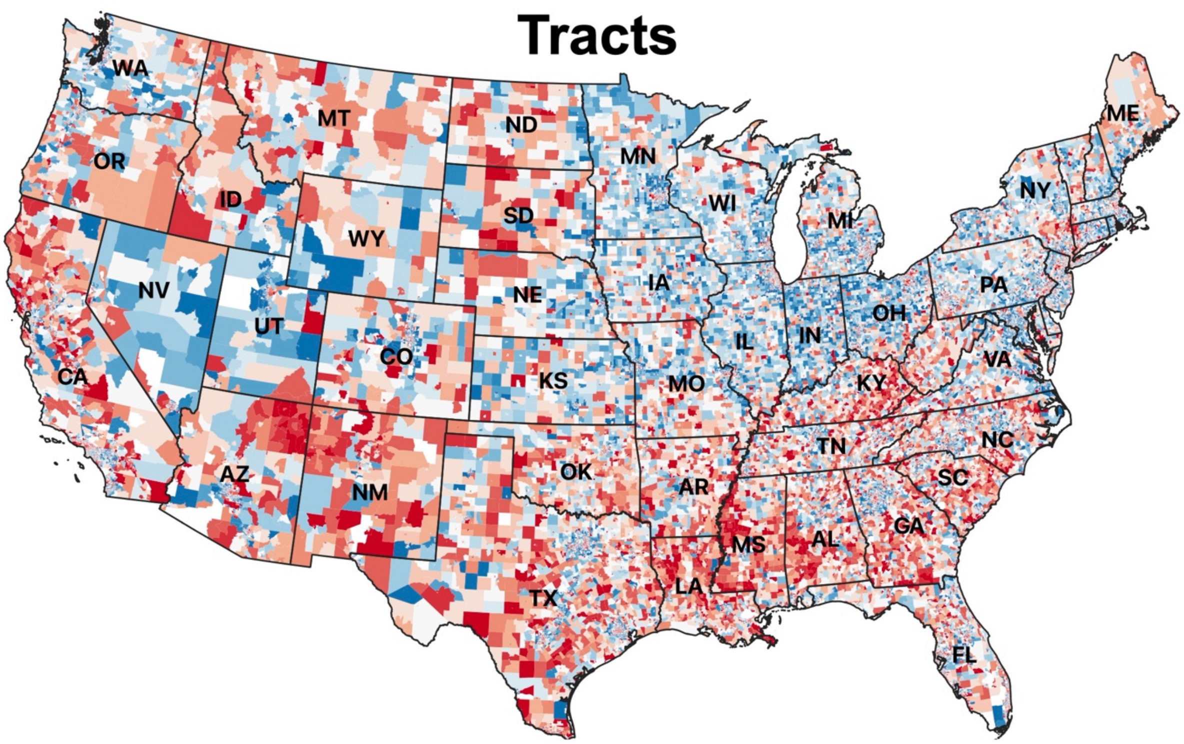

Counties
WA
MT
ND
ME
MN
OR
ID
WI
NY
SD
WY
MI
IA
PA
NV
NE
OH
UT
IL
IN
CO
VA
CA
KS
KY
MO
NC
TN
AZ
OK
AR
SC
NM
GA
MS
AL
LA
TX
FL
Tracts
Percent
0 - 0.025
0.025 - 0.05
0.05 - 0.06
0.06 - 0.07
0.07 - 0.08
0.08 - 0.09
0.09 - 0.1
0.1 - 0.125
0.125 - 0.15
0.15 - 0.175
0.175 - 0.2
0.2 - 0.225
0.225 - 0.25
0.25 - 0.5
0.5 - 0.85

Gini Index

Counties



Tracts

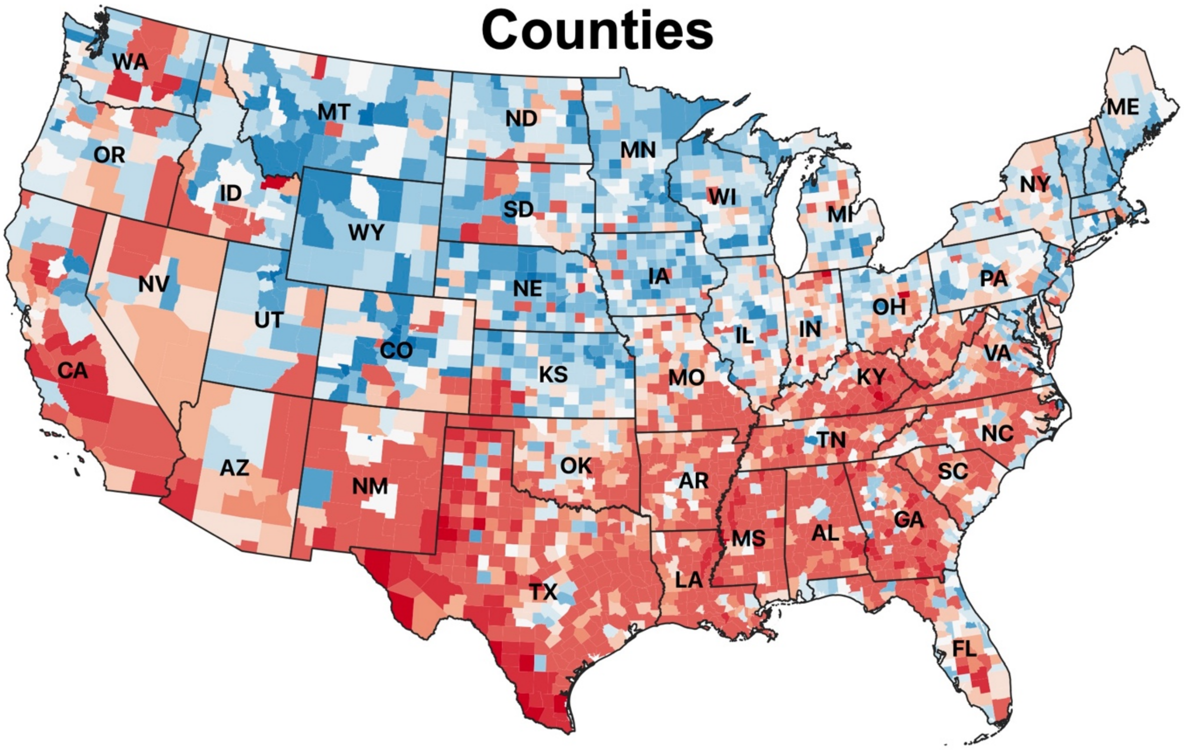


Coefficient (0=perfect equality, everyone receives equal share;
1=perfect inequality, only one recipient/group receives all the income)

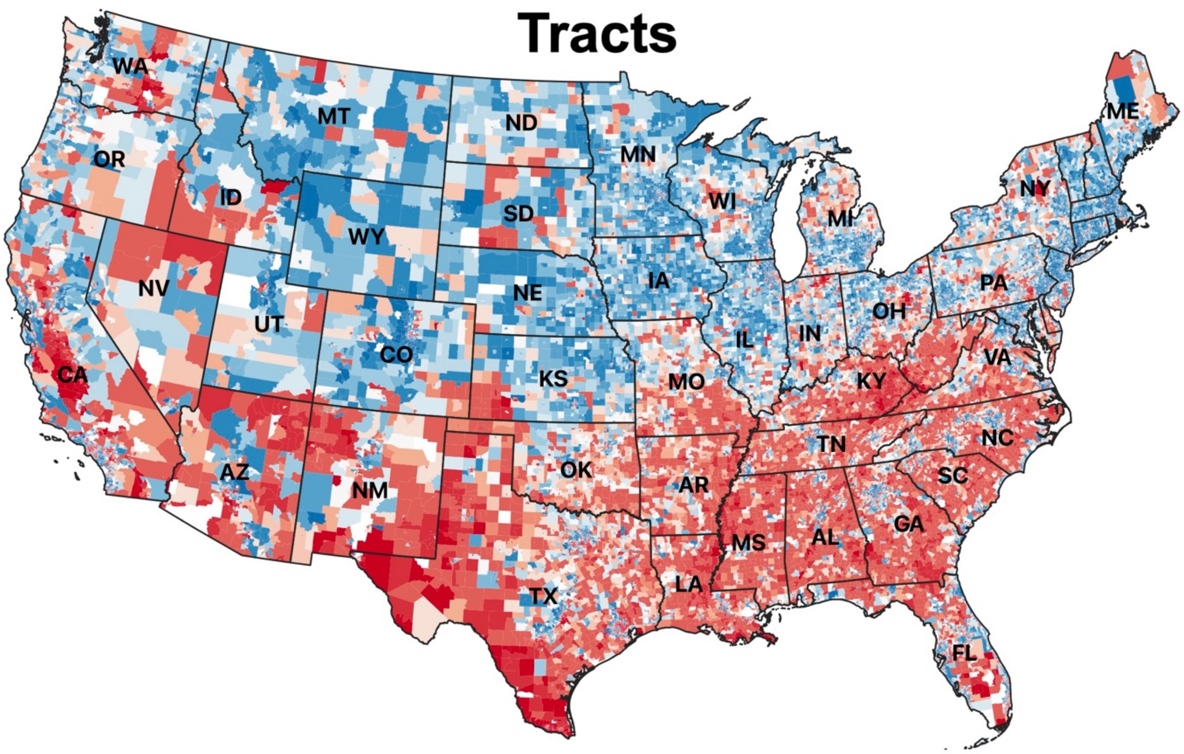
0.05 - 0.35	0.37 - 0.38	0.4 - 0.41	0.44 - 0.45	0.475 - 0.5
0.35 - 0.36	0.38 - 0.39	0.41 - 0.425	0.45 - 0.46	0.5 - 0.525
0.36 - 0.37	0.39 - 0.4	0.425 - 0.44	0.46 - 0.475	0.525 - 0.85

High School Graduates

Counties



Tracts

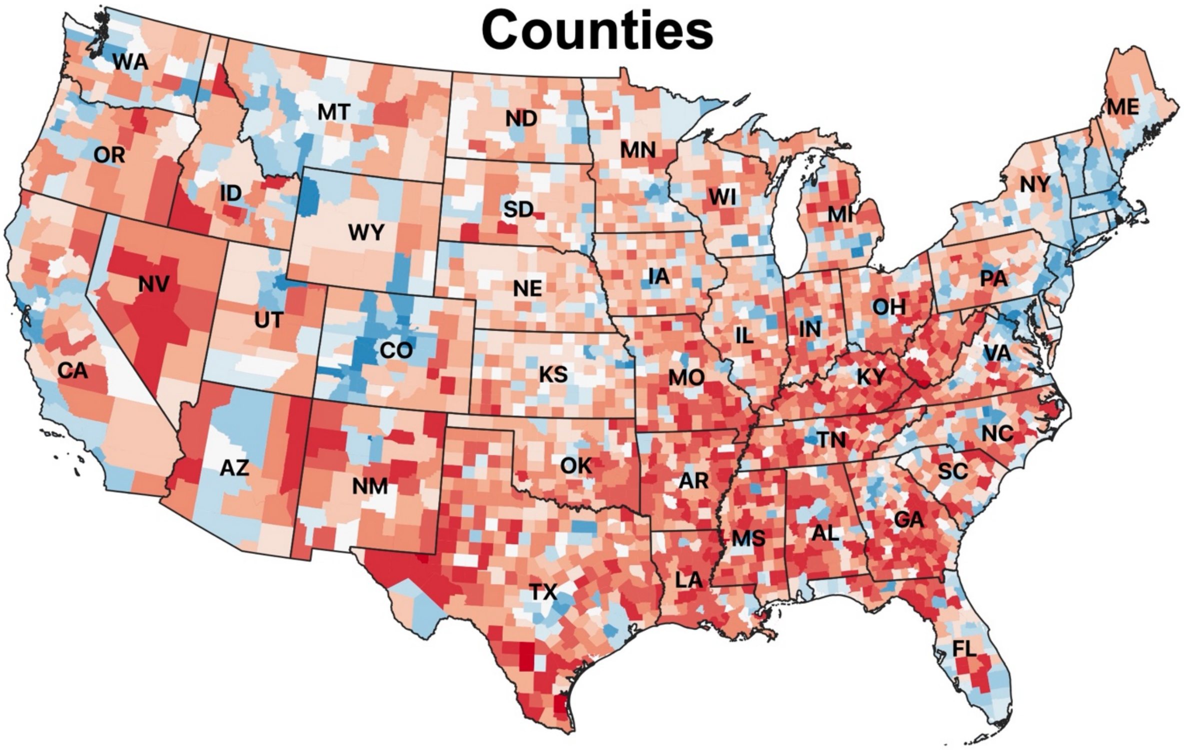


Percent

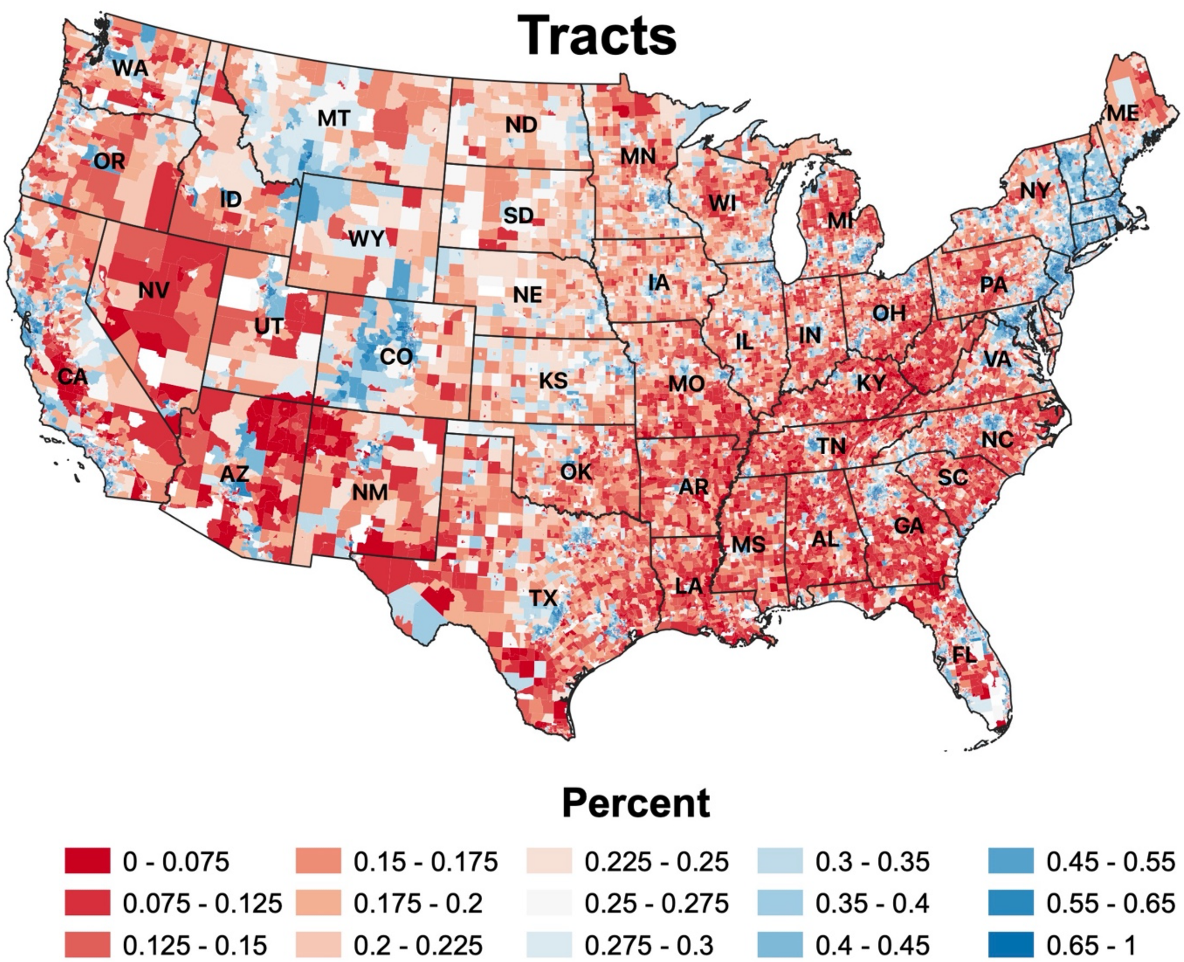
- 0.2 - 0.65
- 0.65 - 0.75
- 0.75 - 0.85
- 0.85 - 0.86
- 0.86 - 0.87
- 0.87 - 0.88
- 0.88 - 0.89
- 0.89 - 0.9
- 0.9 - 0.91
- 0.91 - 0.92
- 0.92 - 0.93
- 0.93 - 0.94
- 0.94 - 0.95
- 0.95 - 0.975
- 0.975 - 1

College Graduates

Counties

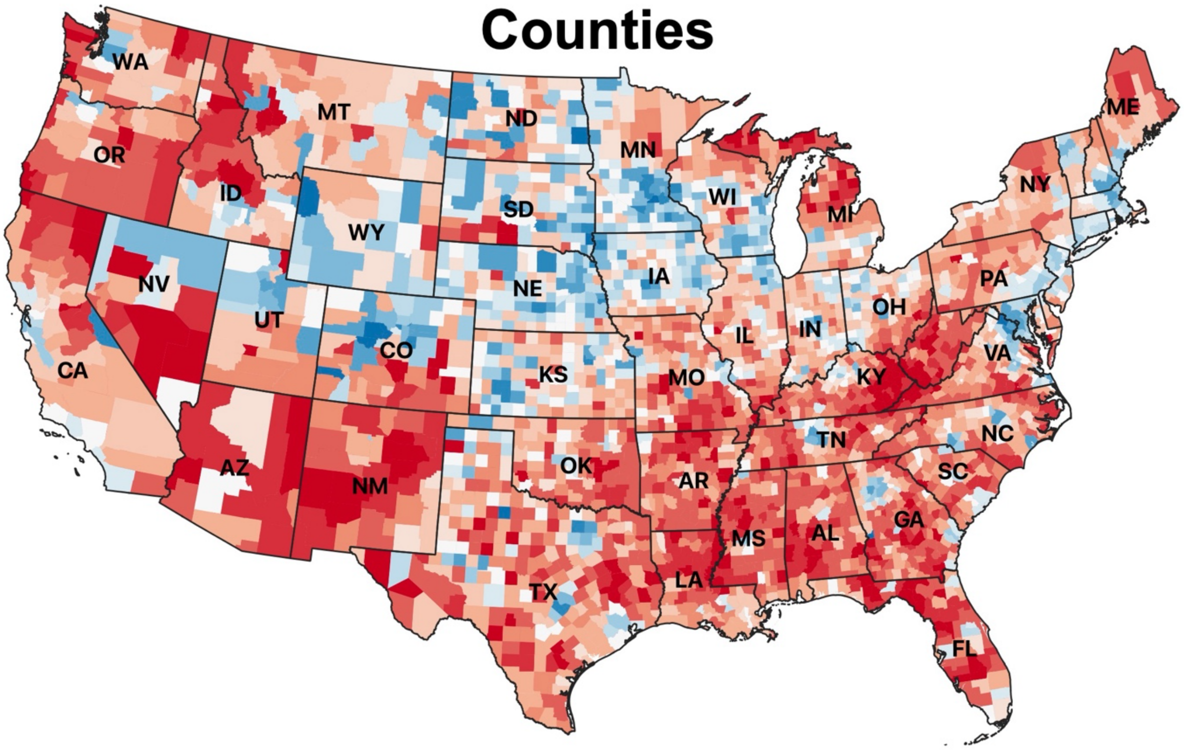


Tracts

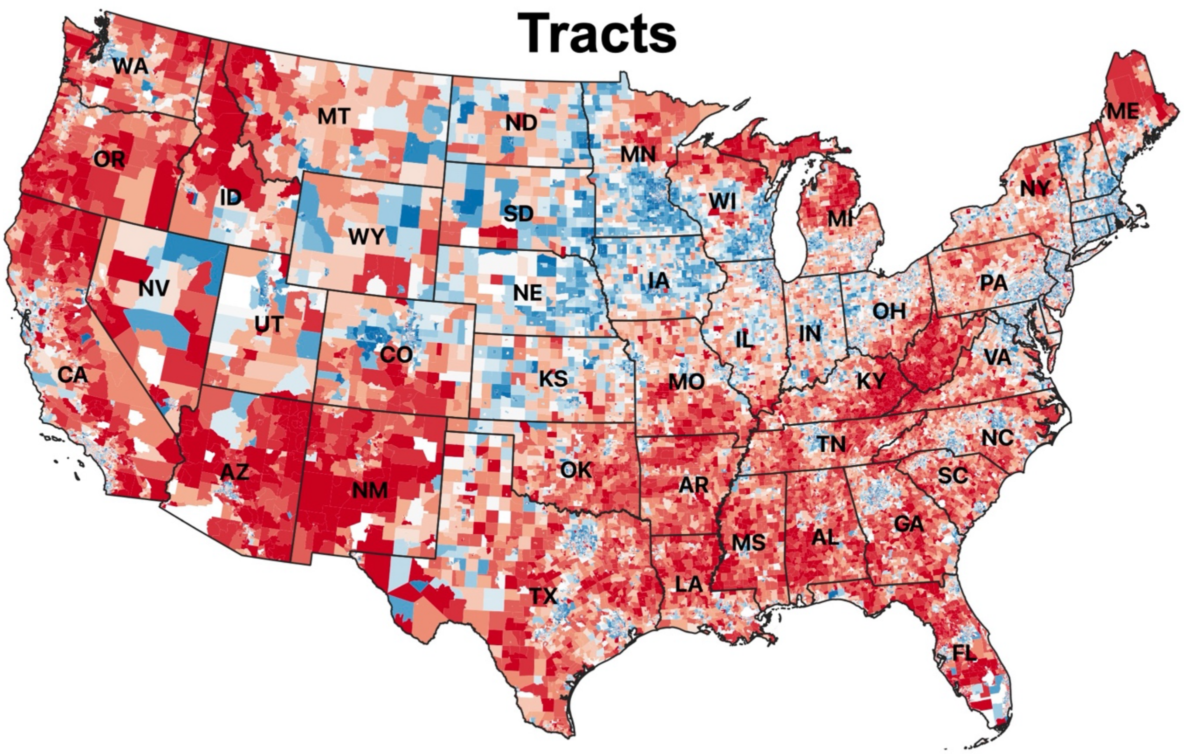


Unemployment

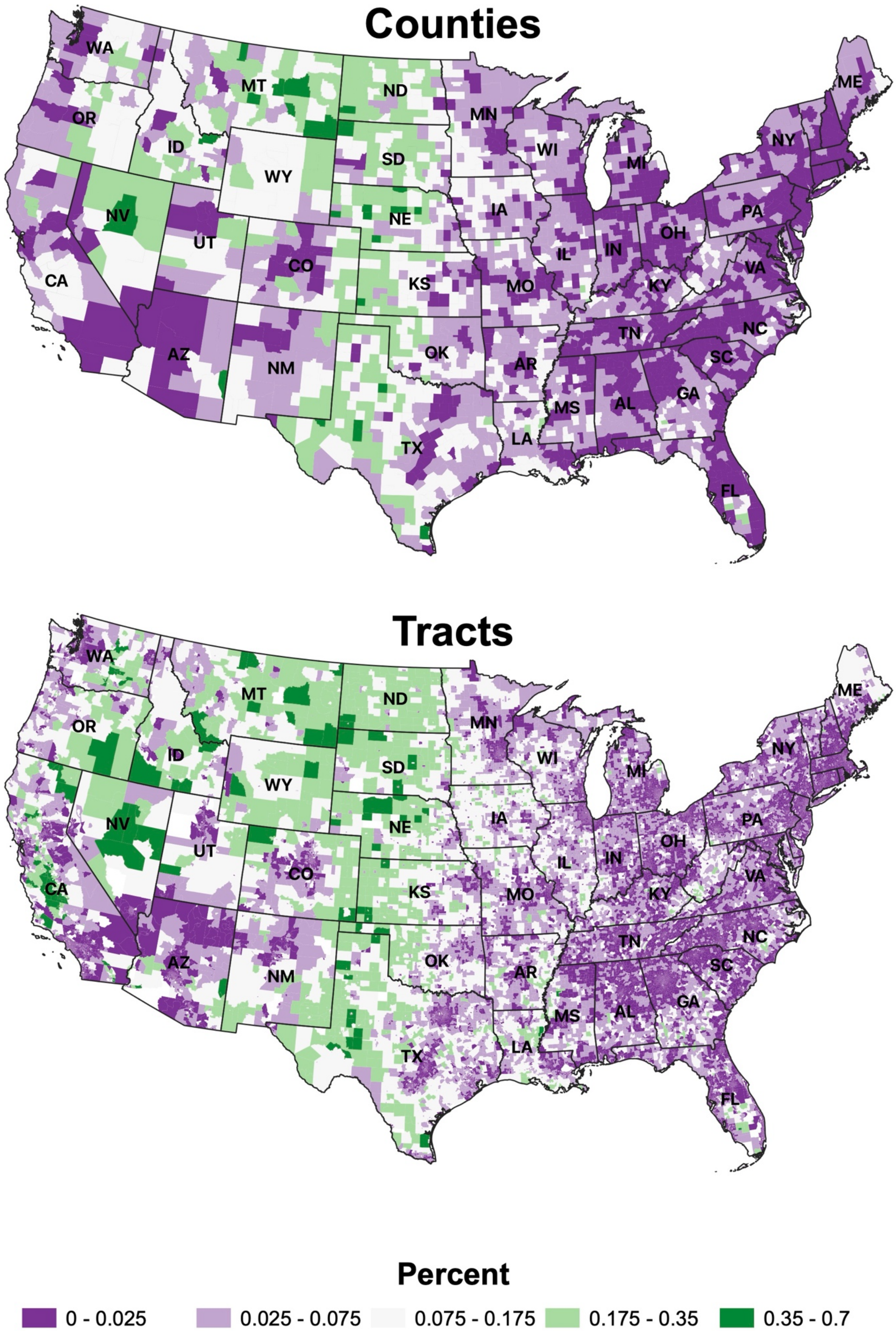
Counties



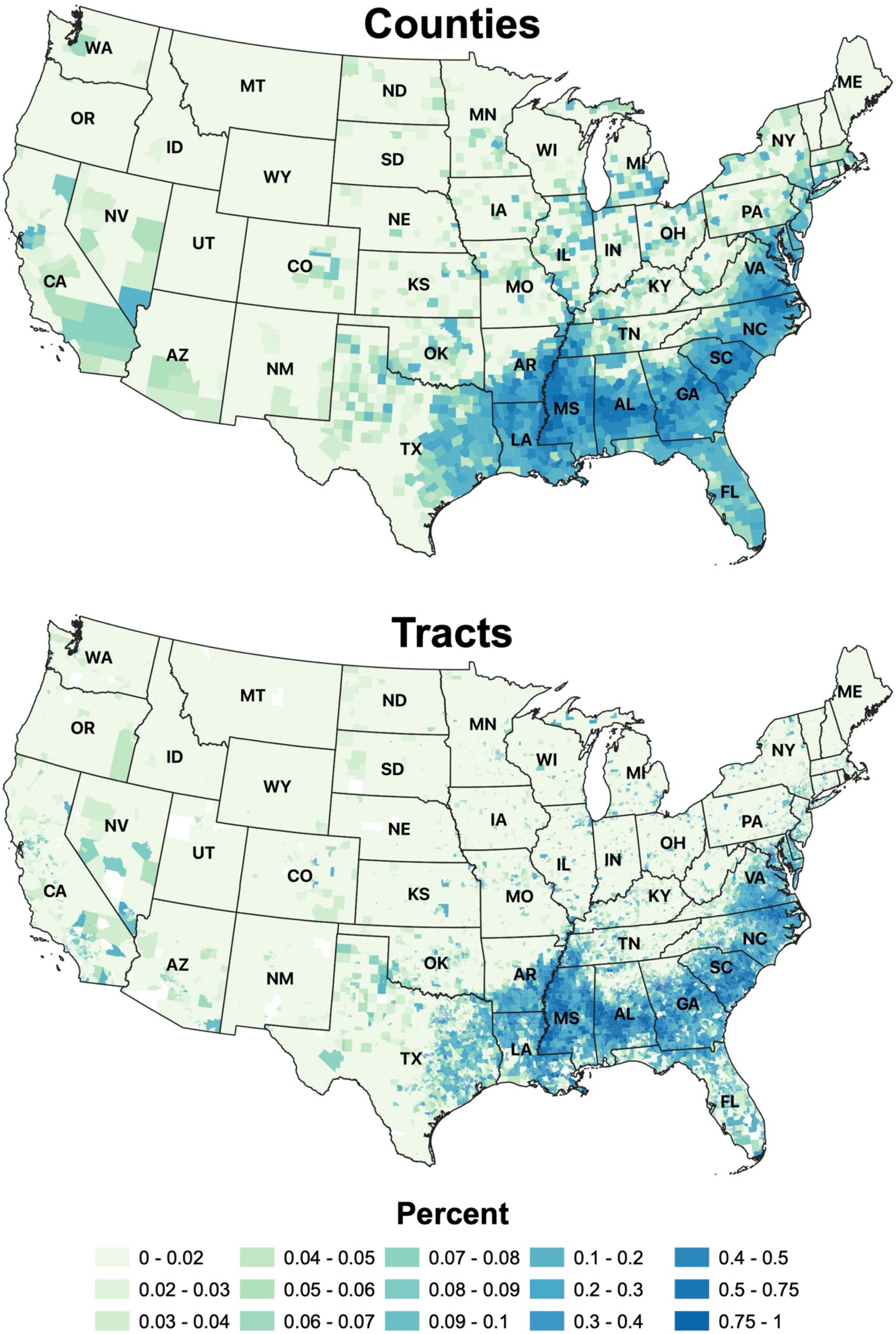
Tracts



Employed Natural Resources

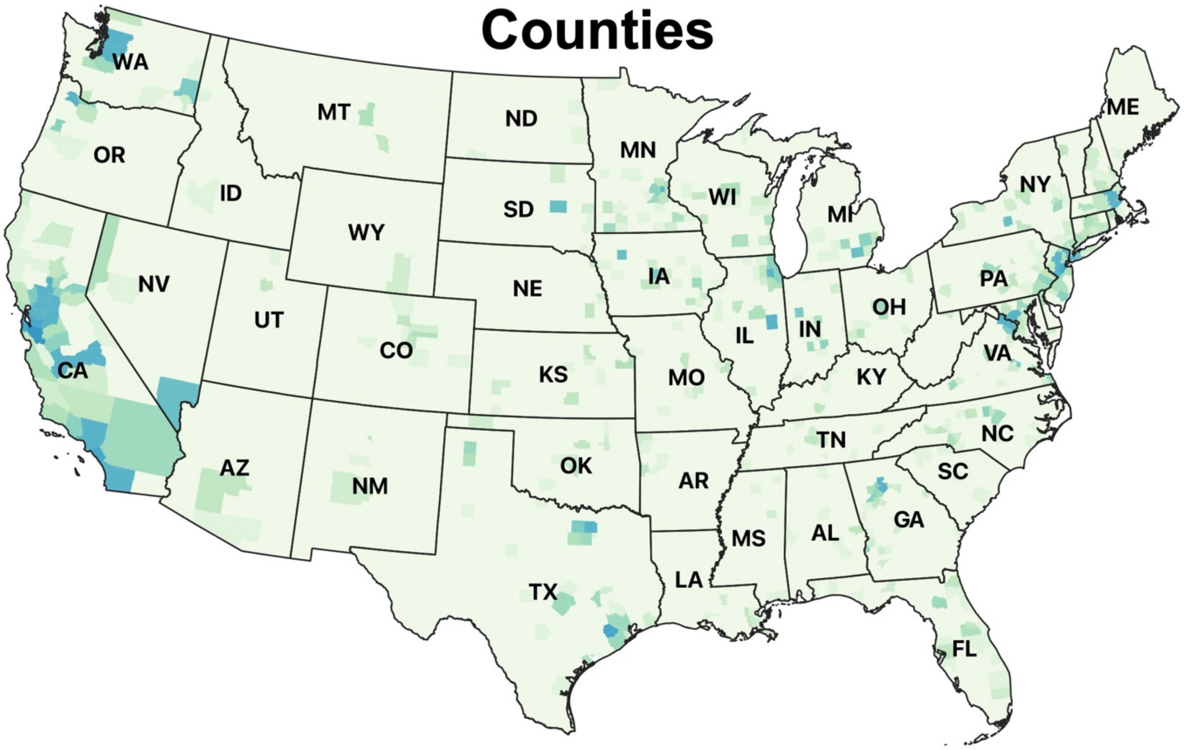


Non-Hispanic Black

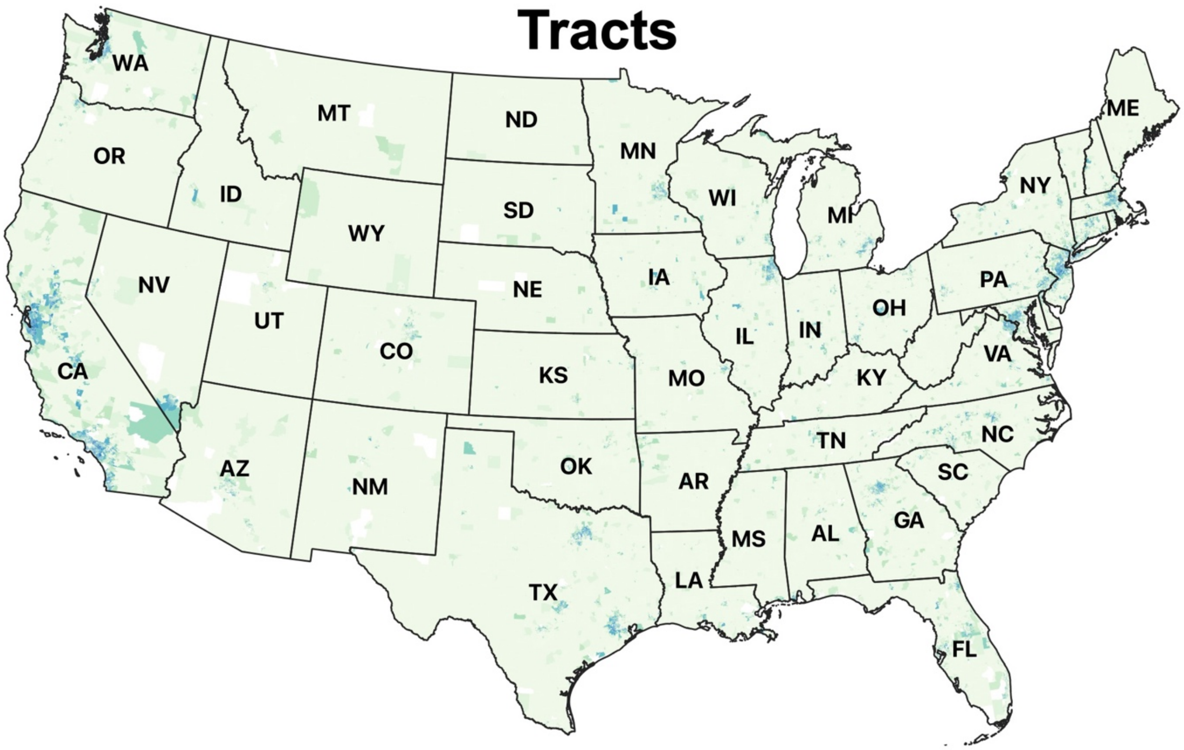


Non-Hispanic Asian

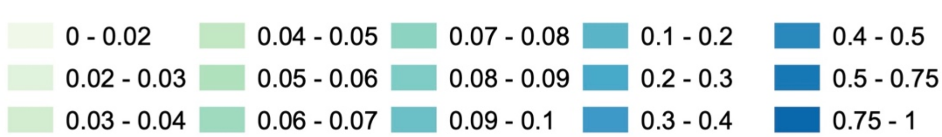
Counties



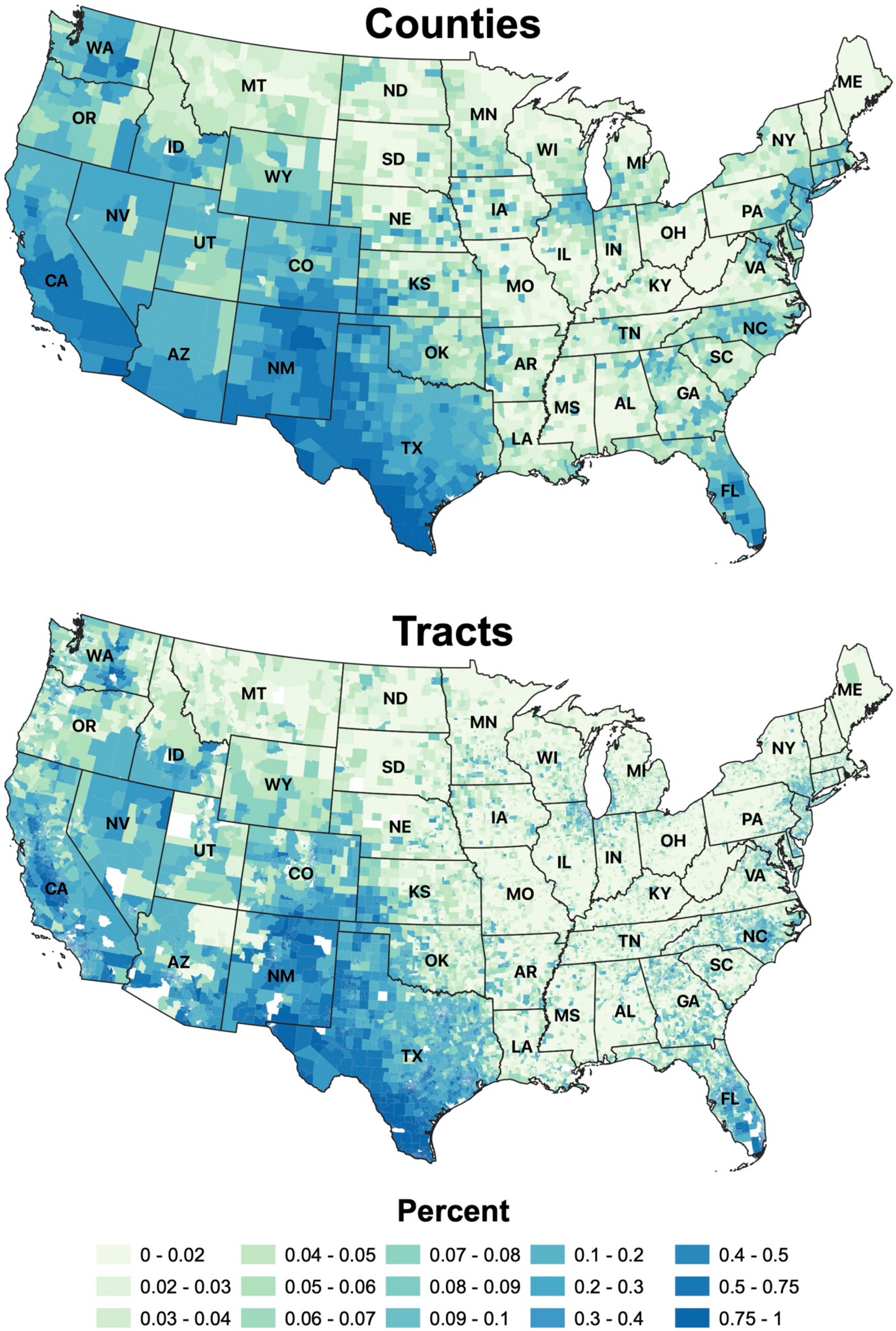
Tracts



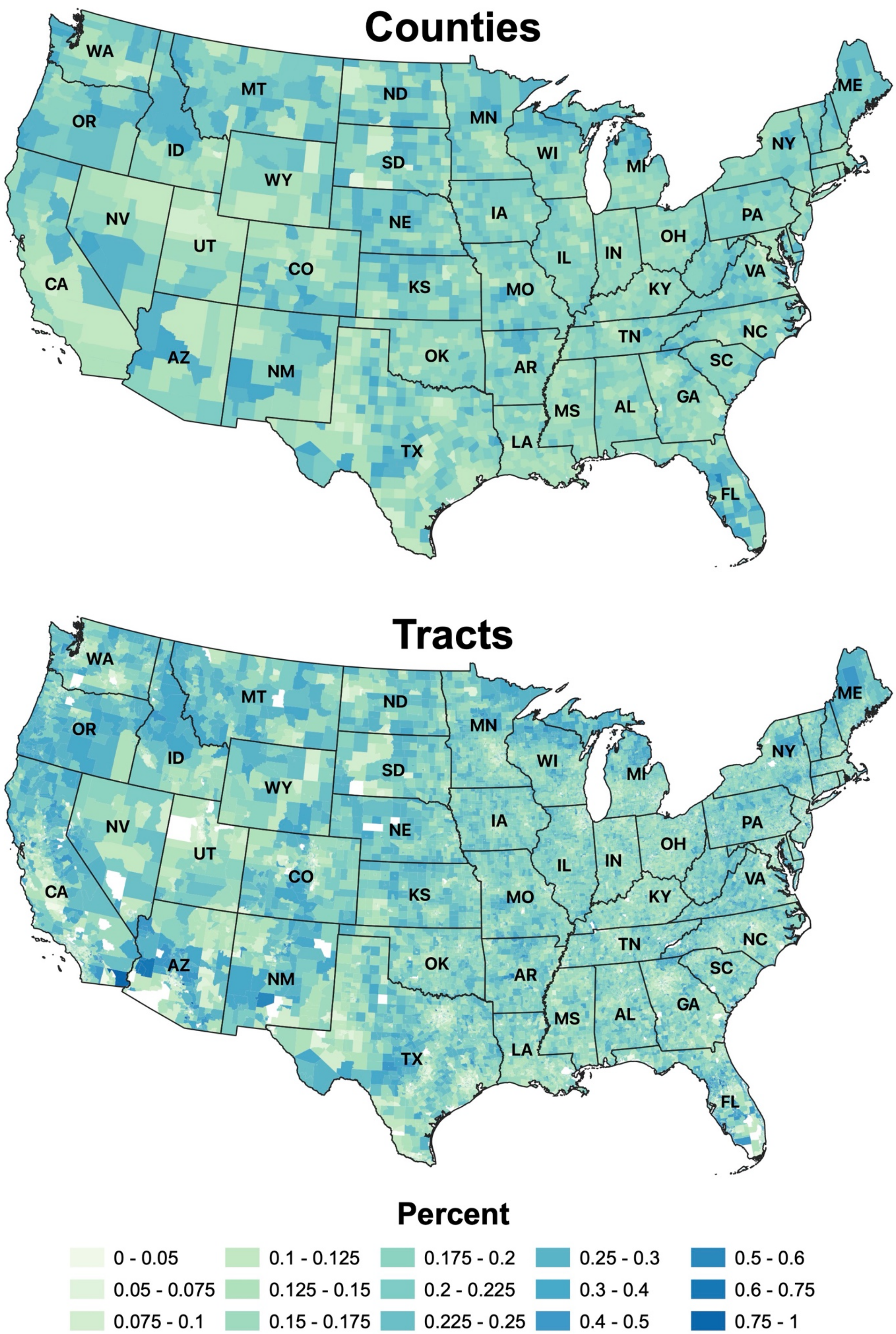
Percent



Hispanic


Counties
WA
MT
ND
ME
OR
MN
ID
SD
WI
NY
MI
WY
NV
NE
IA
PA
UT
OH
IL
IN
CO
VA
CA
KS
MO
KY
NC
TN
OK
AZ
NM
AR
SC
GA
MS
AL
LA
TX
FL
Tracts
WA
MT
ND
ME
OR
MN
ID
SD
WI
NY
MI
WY
NV
NE
IA
PA
UT
OH
IL
IN
CO
VA
CA
KS
MO
KY
NC
TN
OK
AZ
NM
AR
SC
GA
MS
AL
LA
TX
FL
Percent
0 - 0.02
0.02 - 0.03
0.03 - 0.04
0.04 - 0.05
0.05 - 0.06
0.06 - 0.07
0.07 - 0.08
0.08 - 0.09
0.09 - 0.1
0.1 - 0.2
0.2 - 0.3
0.3 - 0.4
0.4 - 0.5
0.5 - 0.75
0.75 - 1

65 Years or Older



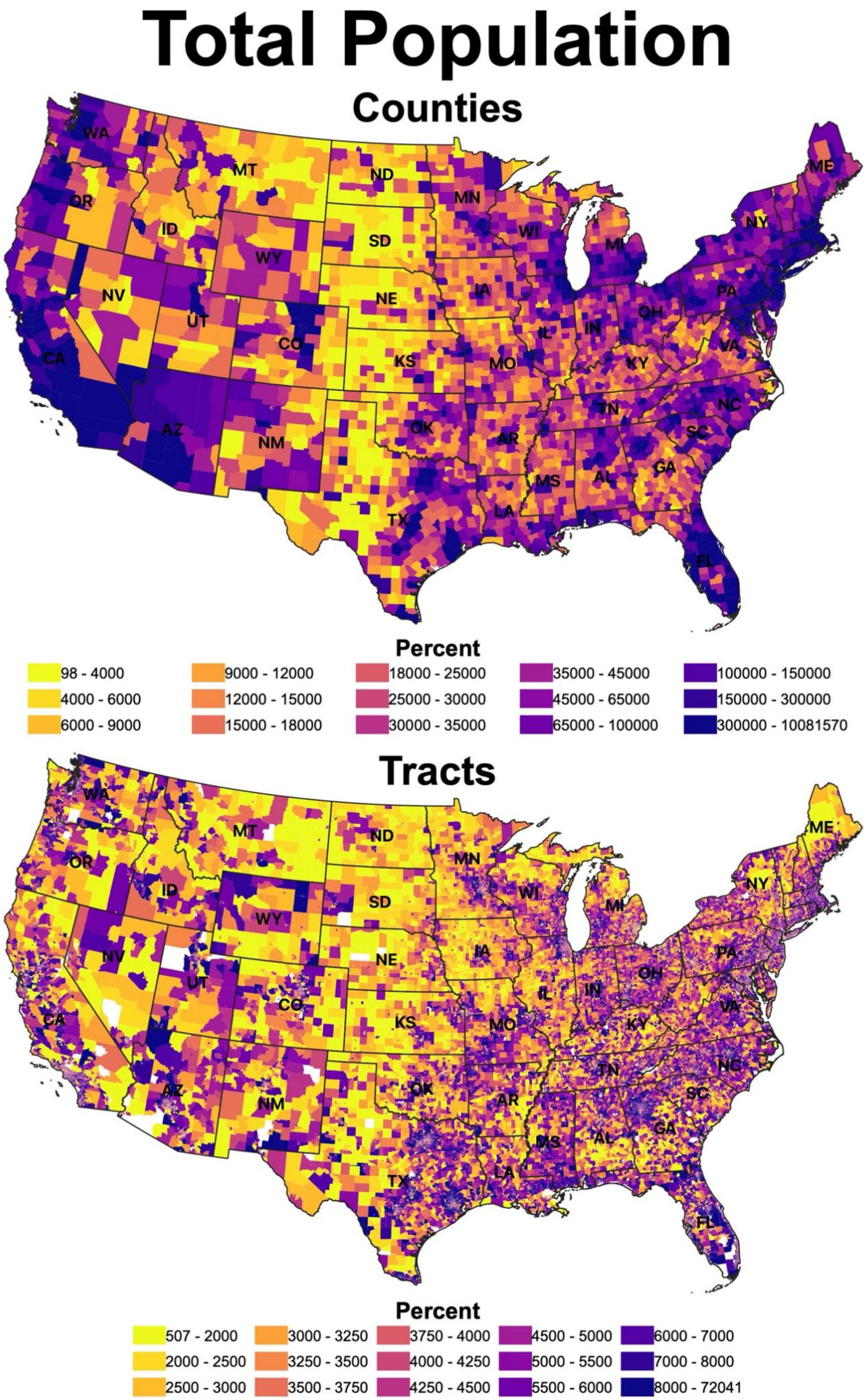


Figure S7. Preceding pages include maps of sociodemographic characteristics considered when validating the new public park cover dataset (PAD-US-AR).

Table S8. Results of generalized linear mixed model regressing sociodemographic characteristics on park cover with state random effects.

	Nationwide				**Northeast**				**Midwest**				**South**				**West**			
	Counties		***Tracts***		***Counties***		***Tracts***		***Counties***		***Tracts***		***Counties***		***Tracts***		***Counties***		***Tracts***	
Predictors	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*
Population density	-0.02	0.18	**-0.02**	**<0.001**	0.08	0.4	**0.09**	**<0.001**	0.05	0.326	**-0.04**	**<0.001**	-0.04	0.221	**-0.07**	**<0.001**	-0.08	0.12	**-0.11**	**<0.001**
Median home value	**0.16**	**<0.001**	**0.05**	**<0.001**	0.05	0.71	0.01	0.669	**0.39**	**<0.001**	**0.03**	**0.012**	**0.25**	**<0.001**	**0.07**	**<0.001**	**0.23**	**0.022**	**0.04**	**0.006**
% poverty	**0.16**	**<0.001**	0	0.582	0.16	0.208	**-0.04**	**0.02**	**0.27**	**<0.001**	**0.05**	**0.002**	**0.18**	**0.001**	0.02	0.073	**0.17**	**0.018**	**-0.04**	**0.002**
Gini index	**-0.09**	**<0.001**	**-0.01**	**0.003**	**-0.33**	**0.008**	-0.01	0.52	**-0.1**	**0.004**	**-0.04**	**<0.001**	**-0.08**	**0.013**	**-0.02**	**0.049**	**-0.15**	**0.009**	0.01	0.237
% high school degree	-0.05	0.056	**0.03**	**<0.001**	**-0.27**	**0.028**	**0.07**	**<0.001**	-0.03	0.536	**0.03**	**0.029**	-0.04	0.438	**-0.03**	**0.006**	-0.11	0.174	-0.03	0.093
% college degree	**0.14**	**<0.001**	**0.06**	**<0.001**	0.32	0.078	**0.04**	**0.007**	0.05	0.347	**0.11**	**<0.001**	-0.06	0.33	0.02	0.19	**0.22**	**0.013**	0.01	0.374
% unemployed	0.04	0.152	**0.09**	**<0.001**	0.1	0.56	**0.1**	**<0.001**	**0.31**	**<0.001**	**0.08**	**<0.001**	0.06	0.226	**0.04**	**<0.001**	-0.16	0.09	**0.09**	**<0.001**
% employed natural resources	**-0.07**	**<0.001**	**0.04**	**<0.001**	**-0.23**	**0.006**	**0.05**	**<0.001**	-0.05	0.195	**-0.08**	**<0.001**	**-0.08**	**0.012**	**-0.02**	**0.002**	-0.05	0.357	**0.1**	**<0.001**
% NH Black	**-0.12**	**<0.001**	**-0.01**	**0.005**	0.05	0.634	-0.02	0.078	-0.02	0.575	0.02	0.123	**-0.27**	**<0.001**	**-0.07**	**<0.001**	-0.05	0.443	**-0.04**	**<0.001**
% NH Asian	**-0.14**	**<0.001**	**-0.03**	**<0.001**	0.07	0.496	**-0.05**	**<0.001**	-0.03	0.425	**0.03**	**<0.001**	0.01	0.853	**0.02**	**0.002**	**-0.29**	**<0.001**	**-0.06**	**<0.001**
% Hispanic	-0.01	0.648	**-0.06**	**<0.001**	0	0.994	0.01	0.37	0.04	0.306	**0.05**	**<0.001**	-0.04	0.417	**-0.06**	**<0.001**	-0.08	0.335	**-0.2**	**<0.001**
% 65+ years	**0.18**	**<0.001**	**0.06**	**<0.001**	**0.38**	**0.001**	**0.08**	**<0.001**	**0.3**	**<0.001**	**0.1**	**<0.001**	**0.19**	**<0.001**	**0.03**	**0.009**	**0.36**	**<0.001**	**0.06**	**<0.001**
% female	**-0.07**	**<0.001**	**-0.05**	**<0.001**	-0.18	0.112	**-0.05**	**<0.001**	**-0.08**	**0.007**	**-0.05**	**<0.001**	-0.04	0.193	**-0.04**	**<0.001**	**-0.17**	**0.001**	**-0.1**	**<0.001**
Total population	**0.06**	**<0.001**	-0.01	0.056	-0.13	0.24	**0.02**	**0.035**	-0.01	0.784	**-0.03**	**<0.001**	**0.15**	**<0.001**	-0.01	0.055	0.09	0.07	**-0.03**	**<0.001**
Random Effects																				
σ^2	0.01		0.01		0.01		0.01		0.01		0.01		0.01		0.01		0.05		0.02	
τ_{00}	0.02 State		0.00 State		0.00 State		0.00 State		0.00 State		0.00 State		0.00 State		0.00 State		0.02 State		0.00 State	
ICC	0.6		0.15		0.07		0.06		0.07		0.04		0.09		0.1		0.26		0.03	
N	49 State		49 State		9 State		9 State		12 State		12 State		17 State		17 State		11 State		11 State	
Observations	3107		70378		217		12882		1055		16751		1421		25579		414		15166	
Marginal R^2 / Conditional R^2	0.064 / 0.628		0.039 / 0.188		0.238 / 0.288		0.038 / 0.092		0.258 / 0.312		0.039 / 0.079		0.150 / 0.227		0.025 / 0.123		0.160 / 0.381		0.146 / 0.176	
AIC	-4095.397		-113179.12		-327.421		-24608.051		-2328.777		-37209.273		-2637.057		-50910.658		-1.875		-13251.152	

Notes: B=standardized betas, σ2 = mean random effect variance, τ_{00} = random intercept variance, ICC = intraclass correlation coefficient, AIC = Akaike Information Criterion

Table S9. Multicollinearity evaluation of generalized linear mixed model regressing socio-demographics on park cover with state random effects at the nationwide scale.

	Variance Inflation Factor (VIF) scores	
	Counties	***Tracts***
Population density	1.4	1.3
Median home value	3.1	2.1
% poverty	3.4	3.4
Gini index	1.8	1.8
% high school degree	3.0	4.1
% college degree	4.0	3.5
% unemployed	4.2	2.7
% employed natural resources	1.3	1.2
% NH Black	1.2	1.5
% NH Asian	2.0	1.3
% Hispanic	1.8	2.4
% 65+ years	2.6	2.7
% female	1.7	1.2
Total population	1.4	1.1

Table S10. Results of generalized linear mixed model regressing sociodemographic characteristics on park cover <u>*in urban areas*</u>.

	Nationwide				Northeast				Midwest				South				West	
	Counties		***Tracts***		***Counties***		***Tracts***		***Counties***		***Tracts***		***Counties***		***Tracts***		***Tracts***	
Predictors	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*
Population density	**-0.15**	**0.011**	**0.02**	**<0.001**	0.04	0.849	**0.11**	**<0.001**	-0.22	0.197	**-0.05**	**<0.001**	**-0.27**	**0.003**	**-0.05**	**<0.001**	**-0.08**	**<0.001**
Median home value	**0.46**	**<0.001**	**0.05**	**<0.001**	**0.58**	**0.049**	-0.02	0.374	0.49	0.068	**-0.09**	**<0.001**	**0.56**	**0.002**	**0.05**	**0.003**	0.03	0.113
% poverty	0.18	0.14	0.01	0.216	-0.01	0.966	0.01	0.718	-0.02	0.95	-0.02	0.384	0.05	0.774	0.01	0.622	0	0.931
Gini index	-0.11	0.236	-0.01	0.21	**-0.57**	**0.031**	0.03	0.051	0.32	0.159	0.02	0.339	0.13	0.363	0	0.85	**-0.03**	**0.013**
% high school degree	0.14	0.219	**0.04**	**<0.001**	-0.19	0.62	**0.11**	**<0.001**	0.19	0.382	0.02	0.453	0.15	0.342	-0.04	0.052	**-0.08**	**0.002**
% college degree	0.08	0.494	**0.15**	**<0.001**	0.11	0.746	**0.19**	**<0.001**	**-0.71**	**0.021**	**0.19**	**<0.001**	-0.12	0.544	**0.17**	**<0.001**	**0.17**	**<0.001**
% unemployed	0.07	0.424	0	0.788	0.14	0.447	**0.07**	**<0.001**	**-0.63**	**0.007**	-0.02	0.261	0.08	0.604	**-0.05**	**0.001**	**-0.04**	**0.013**
% employed natural resources	0	0.942	**-0.02**	**0.005**	-0.26	0.075	-0.02	0.105	**-0.26**	**0.022**	0	0.902	-0.08	0.281	**-0.04**	**<0.001**	-0.02	0.079
% NH Black	0	0.973	**0.05**	**<0.001**	0.42	**0.03**	**0.05**	**<0.001**	**0.44**	**0.013**	**0.05**	**0.005**	0.04	0.638	**0.11**	**<0.001**	**-0.05**	**<0.001**
% NH Asian	**-0.22**	**0.001**	**0.02**	**0.005**	-0.15	0.38	-0.01	0.529	**0.54**	**0.001**	**0.07**	**<0.001**	-0.11	0.253	**0.04**	**<0.001**	0	0.792
% Hispanic	0.17	0.065	**0.04**	**0.001**	0.08	0.741	**0.16**	**<0.001**	-0.03	0.853	-0.01	0.634	**0.29**	**0.028**	**0.08**	**<0.001**	**-0.13**	**<0.001**
% 65+ years	-0.07	0.426	**0.03**	**<0.001**	0.04	0.854	**0.07**	**<0.001**	0.31	0.077	0	0.84	-0.04	0.763	0.02	0.231	0.03	0.084
% female	0.06	0.326	0	0.997	-0.07	0.712	0.02	0.177	-0.1	0.48	0.02	0.156	0.09	0.294	**-0.02**	**0.035**	-0.01	0.322
Total population	**0.25**	**<0.001**	**0.04**	**<0.001**	-0.27	0.098	**0.08**	**<0.001**	0.03	0.841	**0.04**	**0.004**	**0.3**	**0.001**	0	0.922	**0.02**	**0.016**
Random Effects																		
σ^2	0.01		0		0		0		0		0		0		0		0	
τ_{00}	0.01 State		0.00 State		0.00 State		0.00 State		0.00 State		0.00 State		0.00 State		0.00 State		0.00 State	
ICC	0.53		0.11		0.09		0.01		0.33		0.06		0.23		0.2		0.11	
N	40 State		49 State		7 State		9 State		10 State		12 State		15 State		17 State		11 State	
Observations	314		32795		66		7112		69		6656		146		9311		9716	
Marginal R^2 / Conditional R^2	0.159 / 0.601		0.039 / 0.143		0.292 / 0.354		0.076 / 0.085		0.493 / 0.659		0.038 / 0.095		0.308 / 0.464		0.034 / 0.225		0.074 / 0.179	
AIC	-640.7		-89817.576		-153.4		-18169.448		-283.361		-20046.156		-369.313		-28785.3		-24581.512	

Notes: Western counties had too few observations (N=33) to report county-level results, B=standardized betas, σ2 = mean random effect variance, τ_{00} = random intercept variance, ICC = intraclass correlation coefficient, AIC = Akaike Information Criterion