Supplementary Information for: Census-tract aggregation systematically and inequitably understates urban heat exposure across U.S. metropolitan areas

Daniel P. Johnson

Table S1: **Within-tract heterogeneity drives the understatement.** The per-metropolitan understatement rate (top-quintile, $\geq 2\,^{\circ}$C-understated share) regressed across the 54 metropolitan areas by ordinary least squares against three measures of thermal spread, each entered singly. Explanatory power rises monotonically as the predictor is refined toward the variation that averaging removes, the variation *within* tracts, which is the signature of a genuine aggregation mechanism rather than a spurious association with overall temperature range. Corresponds to Fig. 3 of the main text.

Predictor ($n = 54$ metros)	r^2
Metro-wide block-temperature standard deviation	0.08
Mean within-tract temperature standard deviation	0.35
Population-weighted within-tract temperature standard deviation	0.44*

* $p < 0.001$.

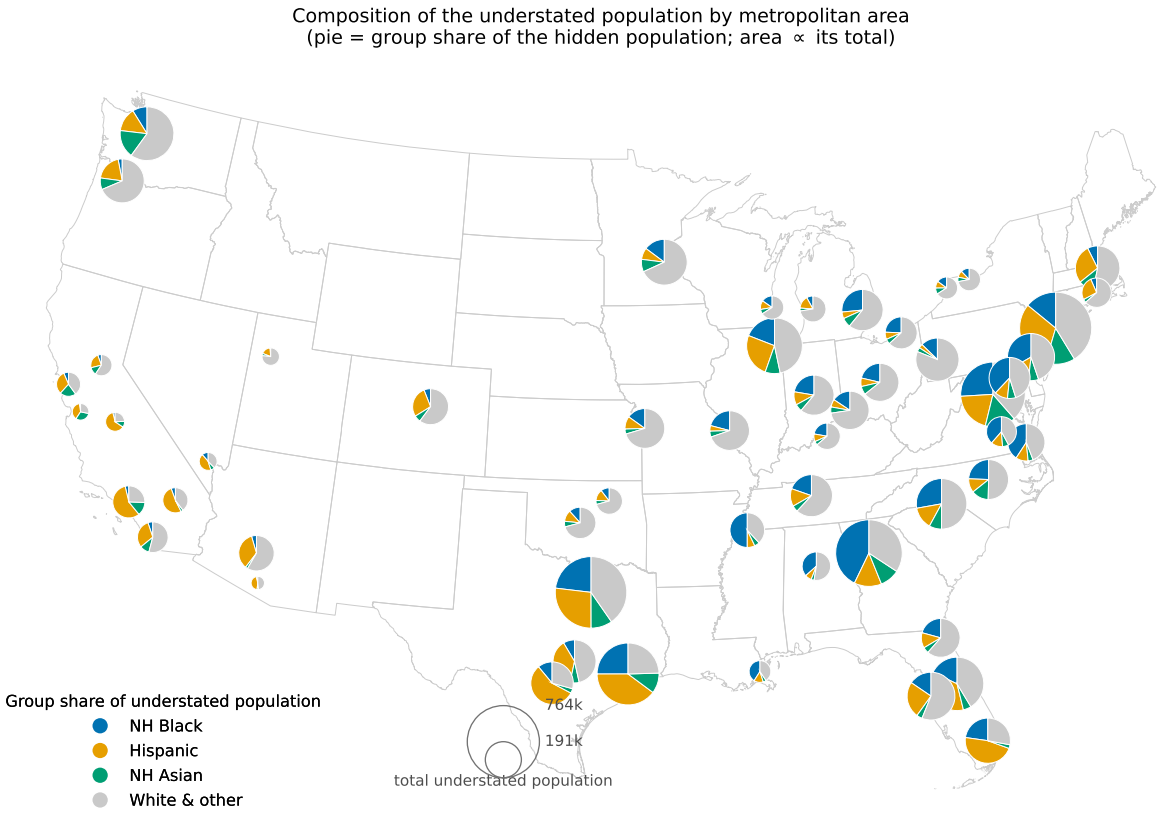

Figure S1: **Composition of the understated population by metropolitan area.** Each metropolitan area is drawn as a pie at its CBSA centroid. Wedges give the racial and ethnic composition of its $\geq 2\,^{\circ}\mathrm{C}$ top-quintile understated population (NH Black, Hispanic, NH Asian, and a combined White-and-other remainder), and pie area is proportional to the total understated population, so visual weight falls on the metropolitan areas where the most residents are hidden by tract averaging. The racial and ethnic makeup of the understated population varies by metropolitan area, as discussed in the main text.

Table S2: **Directional asymmetry of tract misclassification.** Tract averaging misclassifies blocks in both directions, but the two are far from symmetric. The upper panel reports the population misclassified downward (understated: assigned a cooler class than its own block) versus upward (overstated: assigned a hotter class), pooled across the 54 metropolitan areas, at each magnitude threshold and for the extreme tails (top-quintile blocks hidden in lower-class tracts versus bottom-quintile blocks raised to higher-class tracts). The lower panel reports the median across metros of the rate ratio, the share of non-white residents in the misclassified population divided by the same share for white residents, for each set. The understatement is several times larger than the overstatement, and the demographic signal is concentrated in the tails: the hottest hidden blocks are disproportionately minority while the coolest raised blocks are disproportionately white, with the broad misclassifications near-representative. Per-metro values are in the Supplementary Data.

Population misclassified (millions, pooled over 54 metros)			
	Understated	Overstated	Ratio
Any cooler/hotter class, ≥2 °C	37.82	7.56	5.0:1
Any cooler/hotter class, ≥4 °C	16.02	0.69	23.3:1
Extreme tail (top- vs. bottom-quintile)	12.78	3.32	3.9:1
Median rate ratio (non-white vs. white, across 54 metros)			
Set	Median	Metros beyond parity	
Top-quintile understated (hottest hidden)	1.46	50 / 54 above 1	
Bottom-quintile overstated (coolest raised)	0.71	51 / 54 below 1	
Any-cooler understated (broad)	1.02	30 / 54 above 1	
Any-hotter overstated (broad)	1.01	30 / 54 above 1	
Bottom-quintile overstated (coolest raised)	0.86	51 / 54 below 1	

Table S3: **Group-specific hidden-rate ratios at the block-group scale.** For each racial or ethnic group, the rate ratio is the share of that group's residents who fall in the ≥2 °C-understated population, computed with the census block group as the aggregation unit, divided by the same share for non-Hispanic white residents, summarized across the 54 metropolitan areas. A ratio above one indicates that the group is understated more often than white residents. The disparity persists across every group at the block-group scale, complementing the tract-scale equity result in Fig. 4 of the main text and the overall block-group robustness in Fig. S2.

Group	Median rate ratio	Metros above parity
All minority (non-White)	1.36	49 / 54
Non-Hispanic Black	1.38	46 / 54
Hispanic	1.34	49 / 54
Non-Hispanic Asian	1.55	46 / 54

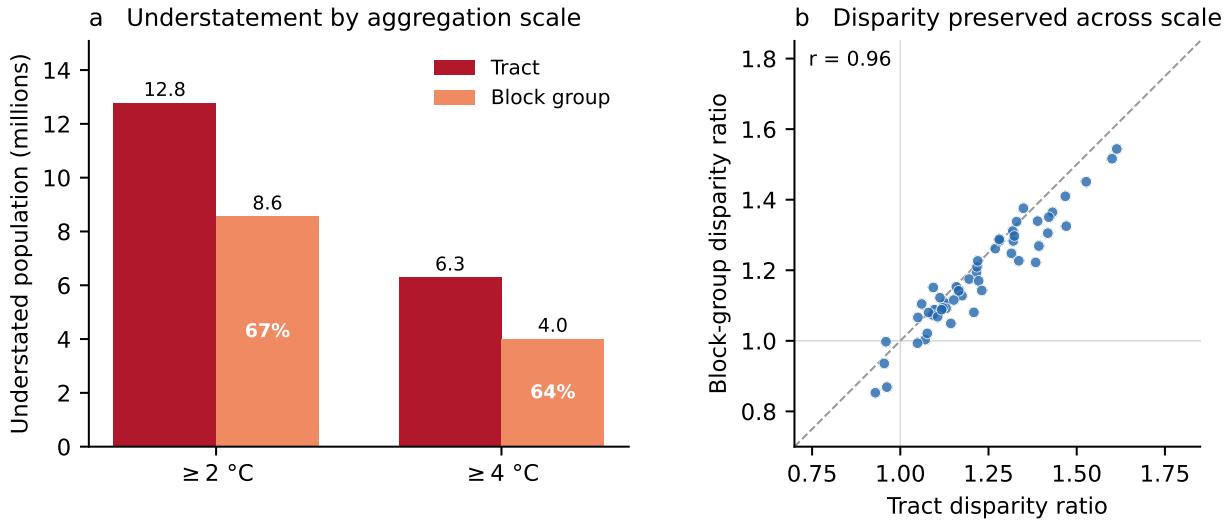

Figure S2: **Robustness to aggregation scale: block group versus tract.** (a) Population in blocks at least 2 °C and at least 4 °C hotter than their unit classification, pooled across the 54 metropolitan areas, under tract and block-group aggregation; the block-group bars retain 67 % and 64 % of the tract-scale understatement respectively. (b) Per-metro hidden-rate ratio (non-white understated rate over white understated rate) under block-group versus tract aggregation, with the 1:1 line; the median falls only from 1.46 at the tract scale to 1.36 at the block group, so the finer unit attenuates the disparity only slightly. The hidden population at block level is zero by construction.

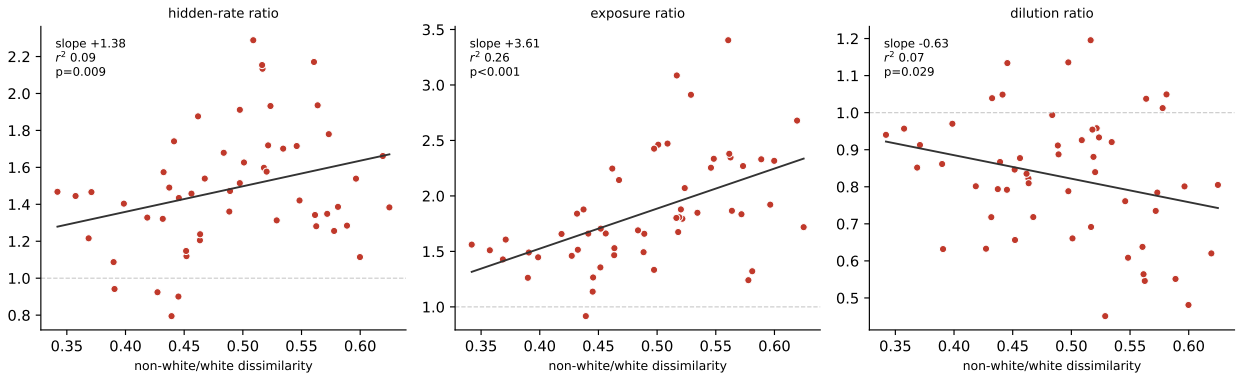

Figure S3: **Decomposition of the disparity against residential segregation.** Each panel regresses a per-metropolitan-area quantity on the non-white/white dissimilarity index across the 54 metropolitan areas, every metro weighted equally, with HC1 robust standard errors. Left: the hidden-rate ratio (non-white understated rate over white understated rate). Center: the exposure ratio (share of non-white residents in hottest-quintile blocks over the white share), which rises steeply with segregation. Right: the dilution ratio (per-block understatement rate among hot-block residents, non-white over white), which declines with segregation, because more segregated metros place minority residents in uniformly hot tracts whose mean already reflects their exposure. Segregation thus acts through two offsetting channels, strong over-exposure and weaker counteracting dilution, so its net effect on the hidden-rate ratio is modest even though both components are individually associated with it; this is why the disparity is only weakly predicted by segregation in the population-weighted national pool.

Table S4: **Sensitivity of the understated population and its rate ratio to analytical choices.** All rows are computed on the 54 metropolitan areas. "Understated, top class" counts residents of top-class blocks whose tract falls in a lower class; "understated, any cooler class" counts residents assigned any cooler class than their own block; both use the 2 °C magnitude guard except in the guard panel. The rate ratio is the median across metros of the rate at which non-white residents fall in the 2 °C top-class understated population (understated non-white residents divided by all non-white residents) relative to the same rate for white residents; values above 1 indicate minority residents are understated more often than white residents. The bias magnitude varies with these choices while the disparity remains above parity throughout.

Configuration	Understated, top class (millions)	Understated, any cooler (millions)	Median rate ratio	Metros > 1 (of 54)
Number of classes (naive zonal tract value, 2 °C guard)				
3 classes	17.70	30.68	1.287	49
4 classes	14.87	35.35	1.408	50
5 classes (primary)	12.78	37.82	1.462	50
6 classes	11.20	39.19	1.497	50
8 classes	9.01	40.62	1.543	50
10 classes	7.57	41.37	1.597	49
Tract value definition (2 °C guard)				
Population-weighted, 3 classes	9.05	12.84	1.162	39
Population-weighted, 4 classes	8.57	15.21	1.320	47
Population-weighted, 5 classes	7.91	16.38	1.425	48
Population-weighted, 6 classes	7.25	17.00	1.493	47
Population-weighted, 8 classes	6.11	17.56	1.586	47
Population-weighted, 10 classes	5.28	17.84	1.612	48
Magnitude guard (quintiles, naive zonal tract value)				
0 °C	20.50	66.44	1.649	53
2 °C	12.78	37.82	1.462	50
4 °C	6.29	16.02	1.275	47

Table S5: **The understatement and its disparity are not artifacts of satellite observation density.** All rows are computed on the 54 metropolitan areas with five exposure classes and the 2 °C magnitude guard. The per-block count of clear daytime Landsat observations was extracted by the same coverage-weighted exact zonal method used for the block temperature, and reproduced the block temperatures to a median absolute difference below 0.001 °C in every metropolitan area. The upper panel retains only blocks meeting an absolute minimum observation count; the lower panel discards the lowest-observed decile and quartile of blocks within each metropolitan area. "Understated" is the top-class population hidden in a lower-class tract; "change" is relative to the unfiltered count; the median and pooled rate ratios compare the understatement rate for non-white residents with that for white residents, with values above 1 indicating minority residents are understated more often. The rate ratio stays between 1.46 and 1.48 (median) across every filter while the understated count falls only in proportion to the blocks removed, indicating the bias does not reside in the sparsely observed blocks.

Observation filter	Understated (millions)	Change (%)	Median rate ratio	Pooled rate ratio	Metros > 1 (of 54)	Blocks removed (%)	Population removed (%)
Absolute minimum clear observations per block							
None (primary)	12.78	0.0	1.46	1.27	50	0.0	0.0
≥ 5	12.78	−0.1	1.46	1.26	50	0.0	0.0
≥ 10	12.76	−0.2	1.47	1.26	50	0.2	0.2
≥ 16	11.53	−9.8	1.48	1.25	50	9.0	8.5
Lowest-observed blocks removed within each metropolitan area							
Lowest 10%	11.48	−10.2	1.47	1.26	51	10.2	9.9
Lowest 25%	9.42	−26.3	1.47	1.24	52	25.2	31.3

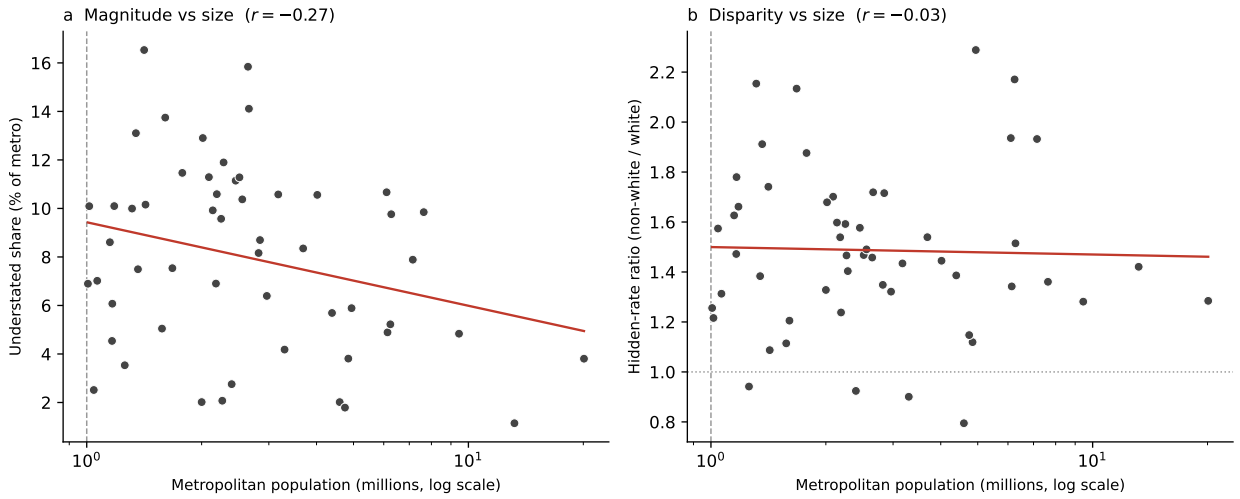

Figure S4: **The understatement and its disparity do not depend on metropolitan size.** (a) Understated share (top-quintile understated population as a percentage of metropolitan population) against metropolitan population on a log scale, with the ordinary-least-squares fit; the association is weakly negative. (b) Per-group hidden-rate ratio against metropolitan population, with the dotted line at parity. The dashed vertical line marks the 1 000 000-population inclusion threshold. Neither quantity rises toward the threshold, so the result is not an artifact of restricting the study to large metropolitan areas and would be expected to hold in smaller ones.

Table S6: **Heat-exposure understatement by census-tract averaging, all 54 metropolitan areas**, ordered by understated share. Rate ratio as defined in Table 1. Complete version of main-text Table 1.

Metropolitan area	Population	Understated ≥2 °C	% of metro	Rate ratio
Raleigh, NC	1,413,982	233,746	16.5	1.74
Orlando, FL	2,645,524	419,104	15.8	1.46
Charlotte, NC–SC	2,660,329	375,399	14.1	1.72
Jacksonville, FL	1,605,789	220,714	13.7	1.21
Memphis, TN–MS–AR	1,345,425	176,288	13.1	1.38

continued on next page

Table S6, continued

Metropolitan area	Population	Understated ≥2 °C	% of metro	Rate ratio
Nashville, TN	2,014,444	259,927	12.9	1.68
Austin, TX	2,283,371	271,646	11.9	1.40
Virginia Beach, VA–NC	1,780,059	204,091	11.5	1.88
Indianapolis, IN	2,089,673	235,944	11.3	1.70
Portland, OR–WA	2,512,859	283,509	11.3	1.47
Pittsburgh, PA	2,455,650	273,689	11.1	1.58
Atlanta, GA	6,104,803	651,218	10.7	1.94
Kansas City, MO–KS	2,192,035	232,111	10.6	1.24
Tampa, FL	3,175,275	335,845	10.6	1.43
Seattle, WA	4,018,762	424,345	10.6	1.44
San Antonio, TX	2,555,942	265,158	10.4	1.49
Oklahoma City, OK	1,425,695	144,808	10.2	1.09
Birmingham, AL	1,180,631	119,212	10.1	1.66
Tulsa, OK	1,015,331	102,487	10.1	1.22
Richmond, VA	1,314,434	131,378	10.0	2.15
Columbus, OH	2,138,926	212,200	9.9	1.60
Dallas, TX	7,637,387	752,096	9.8	1.36
Washington, DC–VA–MD–WV	6,278,536	612,874	9.8	1.51
Cincinnati, OH–KY–IN	2,249,797	215,378	9.6	1.59
Baltimore, MD	2,844,510	247,369	8.7	1.72
Grand Rapids, MI	1,150,015	98,977	8.6	1.63
Minneapolis, MN–WI	3,690,261	308,174	8.4	1.54
St. Louis, MO–IL	2,820,253	230,188	8.2	1.35

continued on next page

Table S6, continued

Metropolitan area	Population	Understated ≥2 °C	% of metro	Rate ratio
Houston, TX	7,149,573	563,905	7.9	1.93
Providence, RI–MA	1,676,579	126,374	7.5	2.13
Louisville/Jefferson County, KY–IN	1,362,180	102,074	7.5	1.91
Rochester, NY	1,065,361	74,759	7.0	1.31
Cleveland, OH	2,180,586	150,517	6.9	1.54
New Orleans, LA	1,007,275	69,472	6.9	1.26
Denver, CO	2,963,821	189,480	6.4	1.32
Buffalo, NY	1,166,902	70,864	6.1	1.78
Boston, MA–NH	4,941,632	291,018	5.9	2.29
Detroit, MI	4,391,947	249,925	5.7	1.39
Philadelphia, PA–NJ–DE–MD	6,245,051	326,185	5.2	2.17
Milwaukee, WI	1,574,731	79,469	5.0	1.11
Miami, FL	6,138,333	300,256	4.9	1.34
Chicago, IL–IN–WI	9,449,351	456,950	4.8	1.28
Fresno, CA	1,164,909	52,862	4.5	1.47
San Diego, CA	3,298,634	137,931	4.2	0.90
Phoenix, AZ	4,845,832	184,611	3.8	1.12
New York, NY–NJ–PA	20,081,935	764,188	3.8	1.28
Salt Lake City, UT	1,257,936	44,471	3.5	0.94
Sacramento, CA	2,397,382	66,074	2.8	0.92
Tucson, AZ	1,043,433	26,243	2.5	1.57
Las Vegas, NV	2,265,461	46,961	2.1	1.47
San Jose, CA	2,000,468	40,354	2.0	1.33

continued on next page

Table S6, continued

Metropolitan area	Population	Understated ≥2 °C	% of metro	Rate ratio
Riverside, CA	4,599,839	92,751	2.0	0.79
San Francisco, CA	4,748,932	84,950	1.8	1.15
Los Angeles, CA	13,200,998	151,132	1.1	1.42
All 54 (pooled)	186,818,779	12,781,651	6.8	1.46 (median)

Figure S5: **The understatement is a scale effect, not an artifact of tract boundaries.** For each of the 54 metropolitan areas, the observed tract-based understated share (marker) is shown against the null distribution of understatement under 100 random, spatially contiguous, size-matched repartitions of the same blocks (grey bar: null mean with 95% interval). The observed value falls at or below its null in all 54 metropolitan areas and significantly below (filled markers; one-sided $p \leq 0.05$) in 46; none lies above its null. Real tract boundaries thus understate no more than, and in most metropolitan areas less than, arbitrary compact partitions of the same blocks, indicating that the bias arises from the scale of aggregation rather than from how the boundaries are drawn.