Reitsma MB, Claypool AL, Vargo J, Shete PB, McCorvie R, Wheeler WH, et al.
Racial/ethnic disparities in COVID-19 exposure risk, testing, and cases at the subcounty level in California.
Health Aff (Millwood). 2021;40(6).

APPENDIX

Geographic Regions

We use the five California Department of Public Health Regions to group PUMAs in some exhibits PUMAs are nested in counties, which are grouped into regions, listed below.

- **North Region**: Butte County, Colusa County, Glenn County, Tehama County, Trinity County, Del Norte County, Lassen County, Modoc County, Plumas County, Siskiyou County, El Dorado County, Humboldt County, Lake County, Mendocino County, Nevada County, Sierra County, Placer County, Sacramento County, Shasta County, Sutter County, Yuba County, Yolo County
- **Bay Area Region**: Alameda County, Contra Costa County, Marin County, Napa County, San Francisco County, San Mateo County, Santa Clara County, Santa Cruz County, Solano County, Sonoma County
- **Central Region**: Alpine County∗, Amador County∗, Calaveras County∗, Fresno County, Inyo County, Kings County, Mariposa County, Mono County, Madera County, Merced County, Monterey County, San Benito County, San Joaquin County, Stanislaus County, Tulare County, Tuolumne County
- **Upper Southern Region**: Kern County, Los Angeles County, San Luis Obispo County, Santa Barbara County, Ventura County
- **Lower Southern Region**: Imperial County, Orange County, Riverside County, San Bernardino County, San Diego County

∗Alpine, Amador, and Calaveras Counties are combined in a single PUMA with Inyo, Mariposa, Mono, and Tuolumne Counties due to small population sizes. California Department of Public Health Regions have Alpine, Amador, and Calaveras Counties included in the North Region, but we include these counties in the Central Region for consistency with PUMA definitions.

Sensitivity Analyses

Distribution of Cases and Tests with Missing Geolocation or Missing Race/Ethnicity

Statewide, there were a total of 817,947 confirmed Covid-19 cases in California between March 22nd and October 3rd, 2020. Of these, 779,823 (95%) had a residential geolocation available that allowed the case to be assigned to a Public Use Microdata Area (PUMA). Over the same time period, there were 15,421,862 tests conducted statewide using polymerase chain reaction (PCR). A smaller proportion of tests (71%, n=10,935,196) had a residential geolocation available, compared to cases. Additionally, 288,376 cases (35% of total cases) and 10,080,062 tests (65.0% of total tests) had unknown/other race/ethnicity. At minimum, all cases and tests had information on county.

In order to provide a realistic estimate of the true level of tests, cases, and test positivity, by race/ethnicity and PUMA, we distributed the tests and cases with unknown residential geolocation and unknown/other race/ethnicity for the main analysis. We took the following approach to distribution:

- Cases and tests with unknown geolocation and known race/ethnicity: use the observed PUMA distribution by race/ethnicity within a county to distribute these cases and tests
- Cases and tests with known geolocation and unknown race/ethnicity: use the observed race/ethnicity distribution within a PUMA to distribute these cases and tests
- Cases and tests with unknown geolocation and unknown race/ethnicity: use the joint observed PUMA and race/ethnicity distribution within a county to distribute these cases and tests

First, we show heatmaps by county showing the percent of tests and cases with residential geolocation information (appendix pages 25-26).

Then, for sensitivity analyses, we present results using distribution based on population size (as opposed to by observed cases and tests) (appendix pages 28-29) and using only data with complete information (appendix pages 30-31). Distribution by population size likely underestimates true disparities, because the assumption is that all unknown cases and tests are distributed proportional to population. Using only data with complete information underestimates true levels of cases and tests, particularly for tests because of the greater proportion of tests with missing information.

Definition of Essential Worker

Six-digit occupation codes are masked to preserve confidentiality for occupations with fewer than 10,000 people nationwide. Full six-digit codes were available for 66% of respondents. Five-digit codes were available for 23%, four-digit codes were available for 10% and three-digit codes were available for the remaining 1% of respondents. We imputed essential worker status based on the weighted-average across jobs within an aggregate category. Ultimately, imputation impacted 15% of individuals in the American Community Survey sample, as the exposure measure was computed at the household level and some aggregate occupational categories were comprised exclusively of either essential or non-essential jobs.

We used the list of jobs classified as essential published by the Department of Homeland Security Cybersecurity and Infrastructure Security Agency (CISA), which has been previously used in multiple studies analyzing essential workers and Covid-19. We tested the sensitivity of our results to more restrictive definitions of essential workers that also include aspects of proximity to other people at work and ability to work from home. These alternative definitions were previously published by Mongey, Pilossoph, and Weinberg in NBER working paper 27085 (https://www.nber.org/papers/w27085). They were constructed using the O NET occupational database. We find that additional restrictions on essential worker based on ability to work from home or proximity to others at work marginally reduced the number of people "exposed" as living in households with fewer rooms than people and at least one essential worker, but did not change the substantive findings, including observed disparities in exposure by race/ethnicity and the observed associations between exposure and outcomes. Using the CISA essential worker classifications alone, approximately 13% of people in California live in a household with fewer rooms than people and at least one essential worker. Further restricting individuals classified as essential workers to those that are unlikely to be able to work remotely, this number drops to 11% of people. Finally, 7% of California's population would be included if we restrict to only those classified as essential workers who are unlikely to be able to work remotely and who are unlikely to be able to practice physical distancing at work.

Although our study did not analyze specific occupations within the broader class of essential workers, we observe differences in the proportional share of essential workers in each 2-digit occupation category.

Exhibit S1: Share of Essential Workers in each Occupation, by Race/Ethnicity

Occupation	All	Asian	Black	Latino	White
Architecture/Engineering	3.3%	7.4%	1.6%	1.2%	4.2%
Building maintenance	5.4%	2.4%	4.6%	9.6%	2.4%
Business/Financial	0.9%	1.1%	1.5%	0.6%	1.2%
Community/Social	1.4%	1.0%	3.4%	1.4%	1.4%
Computer/Math	5.6%	14.1%	3.7%	1.5%	6.7%
Construction/Extraction	7.8%	1.9%	3.7%	12.0%	6.6%
Education	4.2%	2.9%	4.5%	2.6%	6.4%
Entertainment/Media	0.6%	0.3%	0.7%	0.2%	1.1%

Occupation	All	Asian	Black	Latino	White
Farm/Fish/Forest	0.1%	0.0%	0.0%	0.3%	0.1%
Food prep.	7.3%	5.9%	5.1%	10.3%	4.8%
Healthcare supp.	3.0%	3.5%	5.0%	3.4%	1.9%
Healthcare tech.	7.2%	13.3%	10.0%	3.4%	8.4%
Install/Maintenance/Repair	4.4%	2.6%	2.8%	5.0%	4.7%
Legal	0.0%	0.0%	0.0%	0.0%	0.0%
Management	10.9%	11.5%	10.1%	6.1%	16.0%
Material moving	7.0%	3.7%	9.0%	10.5%	4.2%
Office/Admin	13.8%	13.4%	19.1%	12.9%	14.1%
Personal care	0.0%	0.0%	0.0%	0.0%	0.1%
Production	8.2%	8.2%	4.5%	11.9%	4.8%
Protection services	2.2%	0.9%	4.1%	1.8%	2.9%
Sales	4.1%	3.8%	3.1%	3.1%	5.4%
Science	0.9%	1.2%	0.7%	0.4%	1.4%
Transport	1.5%	0.9%	2.8%	1.8%	1.2%

We compared the share of essential workers in each occupational category for Asian, Black, and Latino populations to that of White populations. For example, we observe a greater proportion of Latino essential workers in Maintenance, Construction, Agriculture, Food Preparation, Healthcare Support, Material Moving, Production, and Transport, compared to White essential workers. Differences in exposure, availability of personal protective equipment, paid sick-leave, and other factors affecting risk may vary by occupation within the broad category of essential workers, further contributing to disparities in COVID-19 outcomes.

Exhibit S2: Ratio of Share of Essential Workers in each Occupation to White Population Share

Occupation	Asian	Black	Latino
Architecture/Engineering	1.8	0.4	0.3
Building maintenance	1.0	2.0	4.1
Business/Financial	0.9	1.2	0.5
Community/Social	0.7	2.4	0.9
Computer/Math	2.1	0.6	0.2
Construction/Extraction	0.3	0.6	1.8
Education	0.4	0.7	0.4
Entertainment/Media	0.3	0.6	0.2
Farm/Fish/Forest	0.1	0.2	3.9
Food prep.	1.2	1.1	2.1
Healthcare supp.	1.8	2.6	1.8
Healthcare tech.	1.6	1.2	0.4
Install/Maintenance/Repair	0.6	0.6	1.1
Management	0.7	0.6	0.4
Material moving	0.9	2.1	2.5

Occupation	Asian	Black	Latino
Office/Admin	1.0	1.4	0.9
Personal care	0.4	0.7	0.7
Production	1.7	0.9	2.5
Protection services	0.3	1.4	0.6
Sales	0.7	0.6	0.6
Science	0.8	0.5	0.3
Transport	0.7	2.2	1.4

Temporal Changes in Exposure from the American Community Survey

The most recent 5-year microdata available from the American Community Survey cover the time period from 2014-2018. Although there is certainly dynamic change in occupation and housing during the five year time period and also in the post-survey period since 2018, at the PUMA-level, results are relatively constant, mitigating some concern about this potential data limitation. The rank correlation comparing PUMAs in 2014 to PUMAs in 2018 is 0.84. We present annual estimates from 2014-2018 by PUMA (appendix page 8) and by race/ethnicity within each PUMA (appendix page 9).

Exhibit S3. Sensitivity analysis for time trends in ACS joint-exposure indicator. Lines are PUMAs, darker shade indicates larger population.

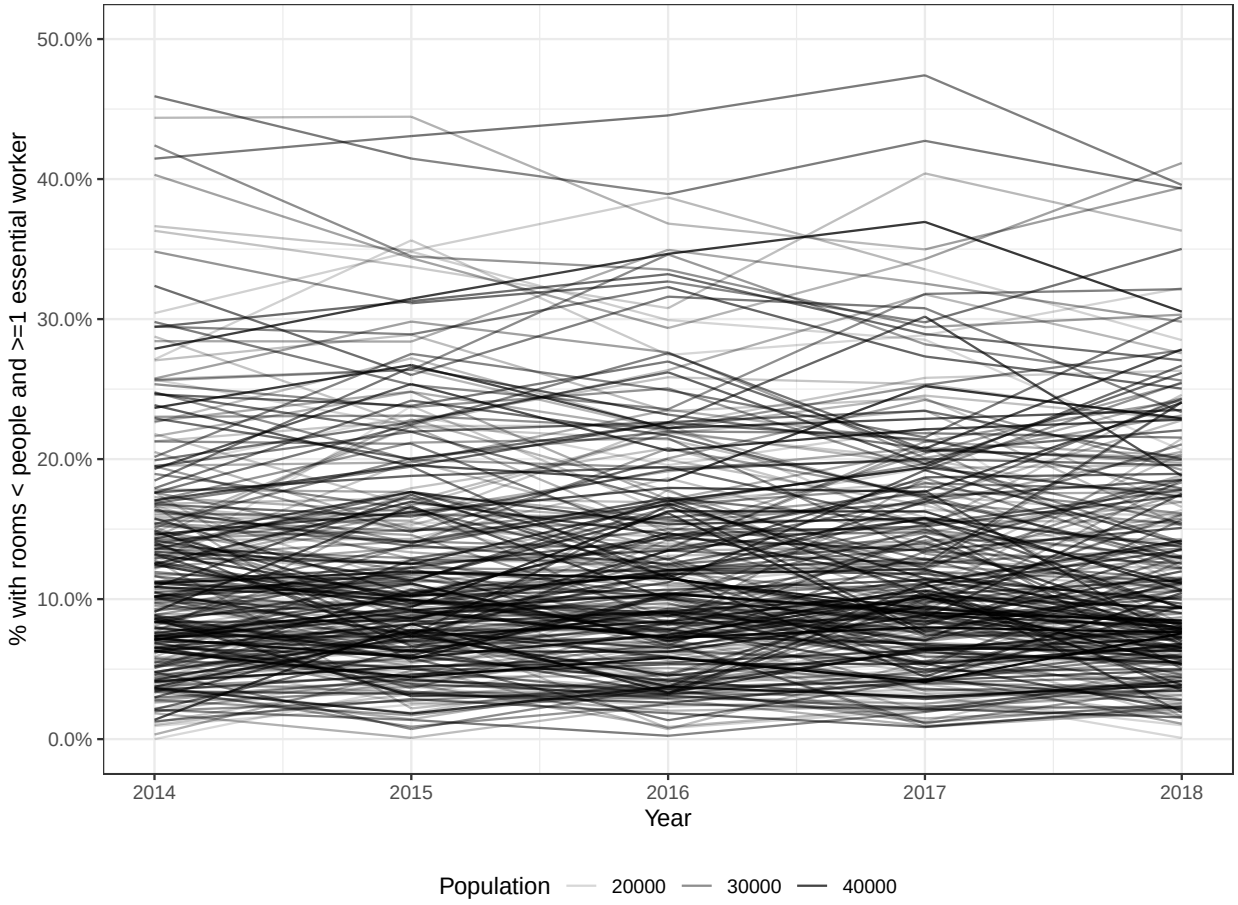
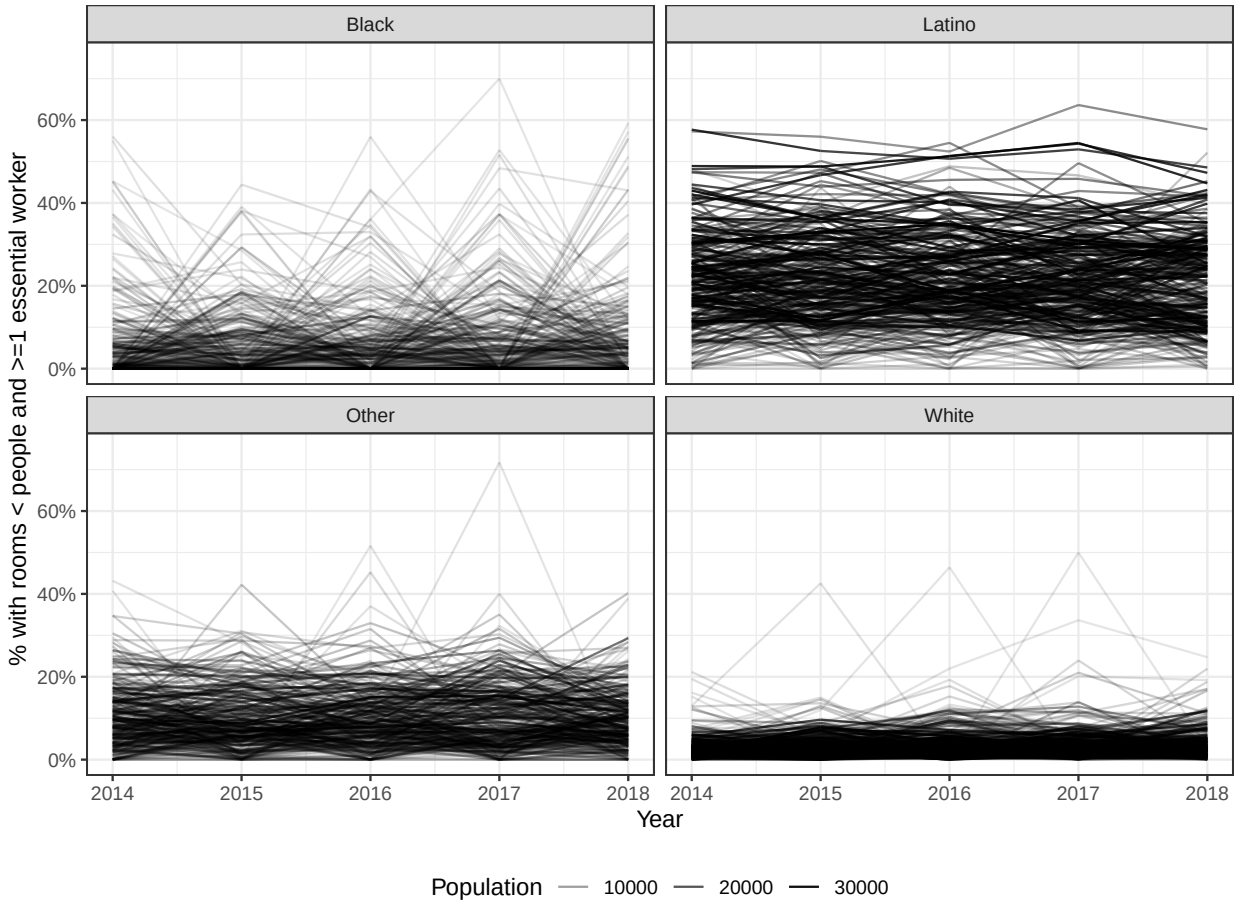


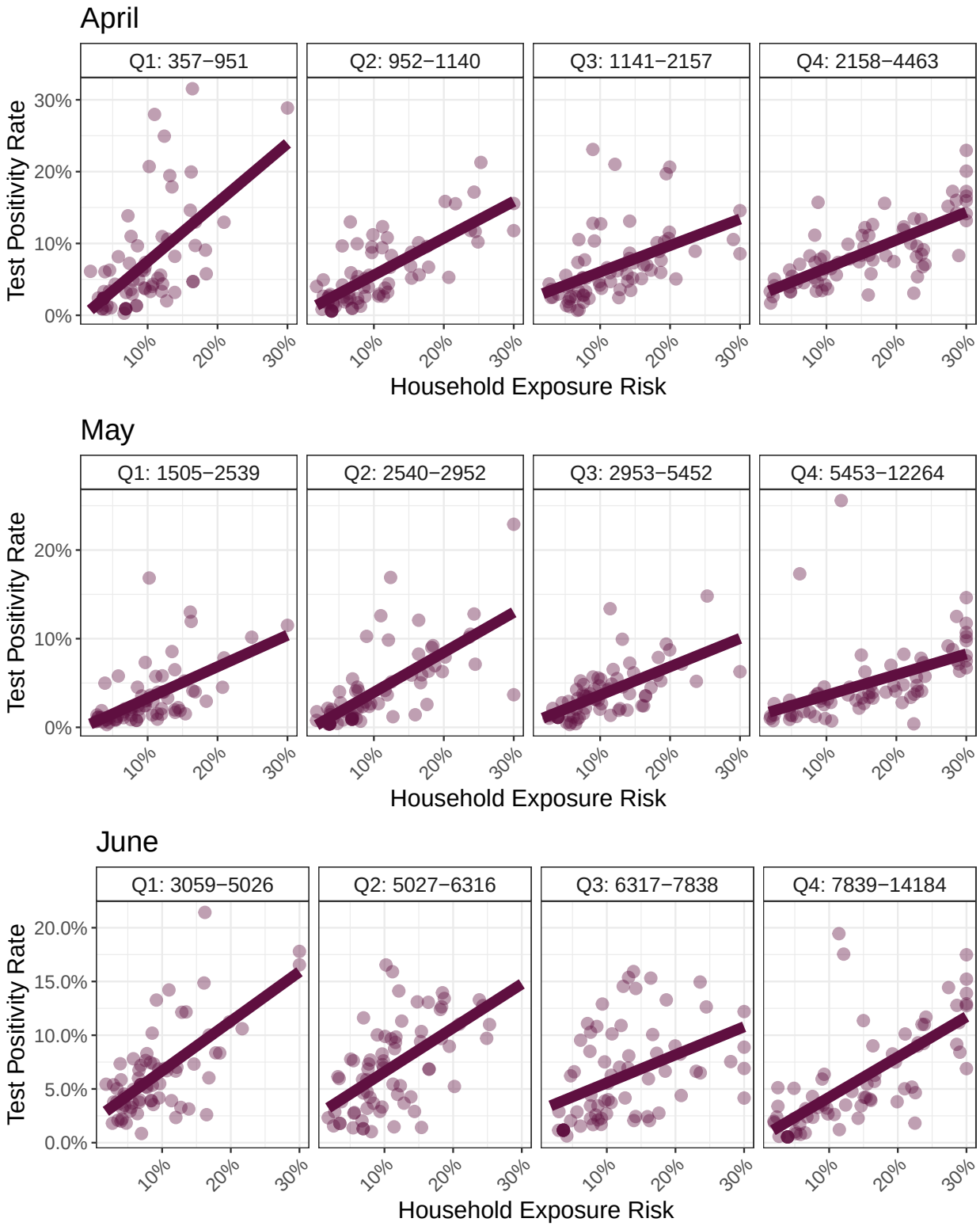
Exhibit S4. Sensitivity analysis for time trends in ACS joint-exposure indicator, by race/ethnicity. Lines are PUMAs, darker shade indicates larger population. Some PUMAs have unstable estimates by race/ethnicity due to small populations, motivating the use of the five-year combined estimates since there are negligible temporal trends.



Supplemental Results

Here we present manuscript Exhibits 3 and 4 for all months April through September and separate by region (appendix pages 11-23). We also show scatter plots between test rate and test positivity, by race/ethnicity, region, and month. Additionally, we show scatter plots between our joint measure of household exposure risk and case rates and test rates. Each point represents one PUMA, and results are shown by month and region (appendix pages 32-33). Finally, we present maps of plurality race/ethnicity, exposure risk, test rate, and case rate, for each of the five regions (appendix pages 34-38) and for counties with more than four PUMAs (appendix pages 39-52). Scales for the plots vary to highlight within-county differences.

Exhibit S5. Household exposure risk vs. test positivity rate, by PUMA and test rate quartile, shown for for April-September for all PUMAs in California.



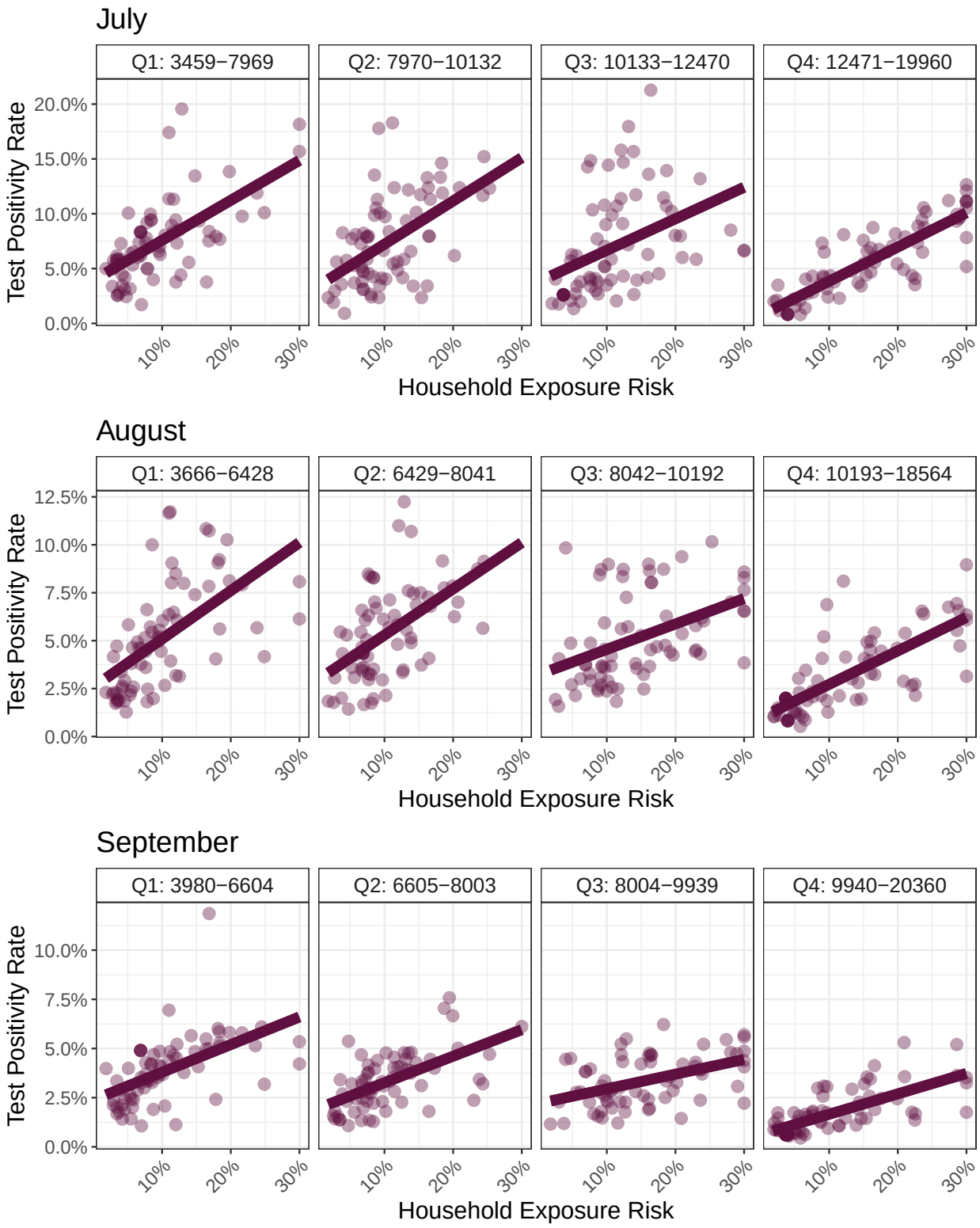
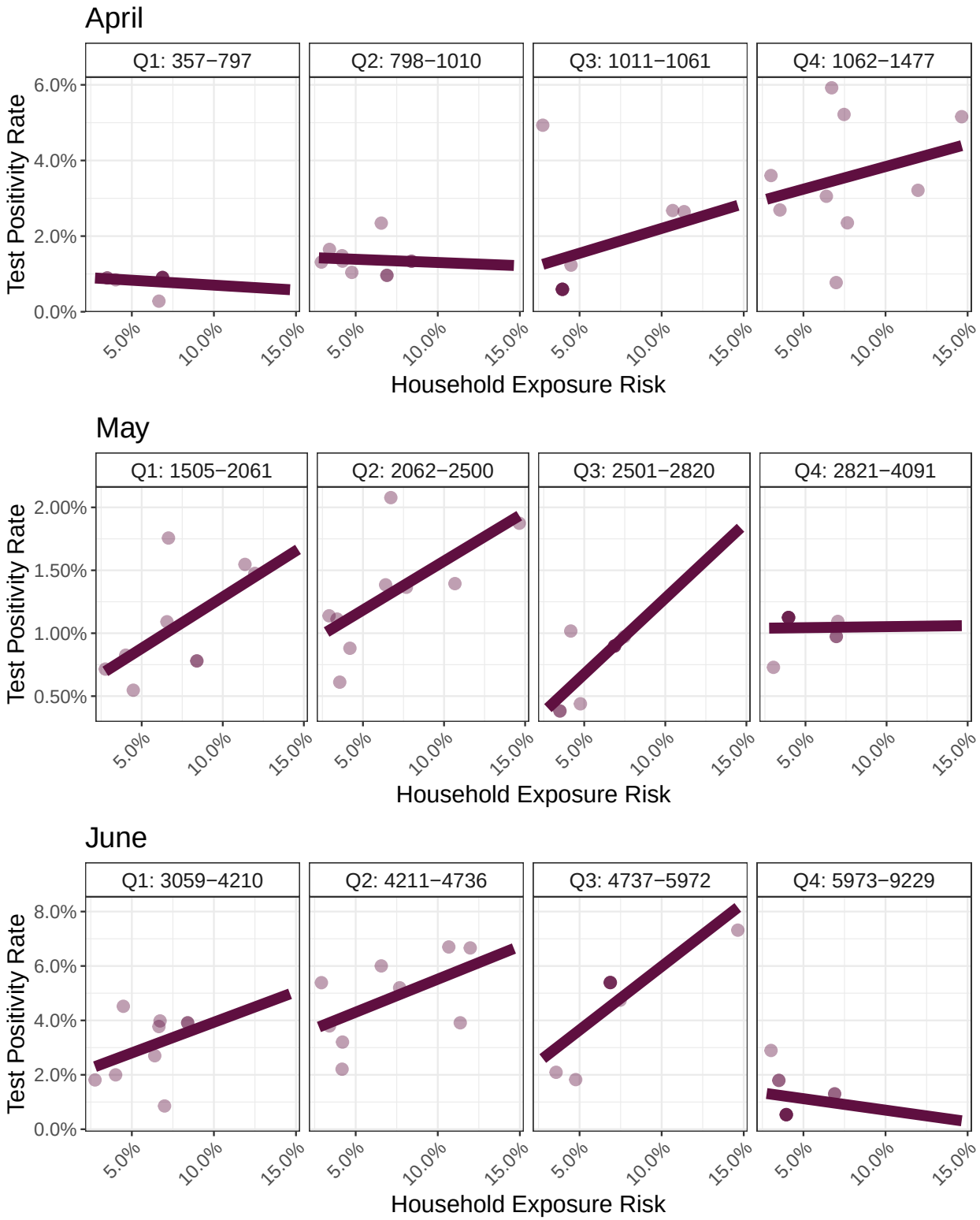

July
Q1: 3459−7969
Q2: 7970−10132
Q3: 10133−12470
Q4: 12471−19960
Test Positivity Rate
20.0%
15.0%
10.0%
5.0%
0.0%
10%
20%
30%
Household Exposure Risk
August
Q1: 3666−6428
Q2: 6429−8041
Q3: 8042−10192
Q4: 10193−18564
Test Positivity Rate
12.5%
10.0%
7.5%
5.0%
2.5%
0.0%
10%
20%
30%
Household Exposure Risk
September
Q1: 3980−6604
Q2: 6605−8003
Q3: 8004−9939
Q4: 9940−20360
Test Positivity Rate
10.0%
7.5%
5.0%
2.5%
0.0%
10%
20%
30%
Household Exposure Risk

Exhibit S6. Household exposure risk vs. test positivity rate, by PUMA and test rate quartile, shown for for April-September for all PUMAs in the North Region.



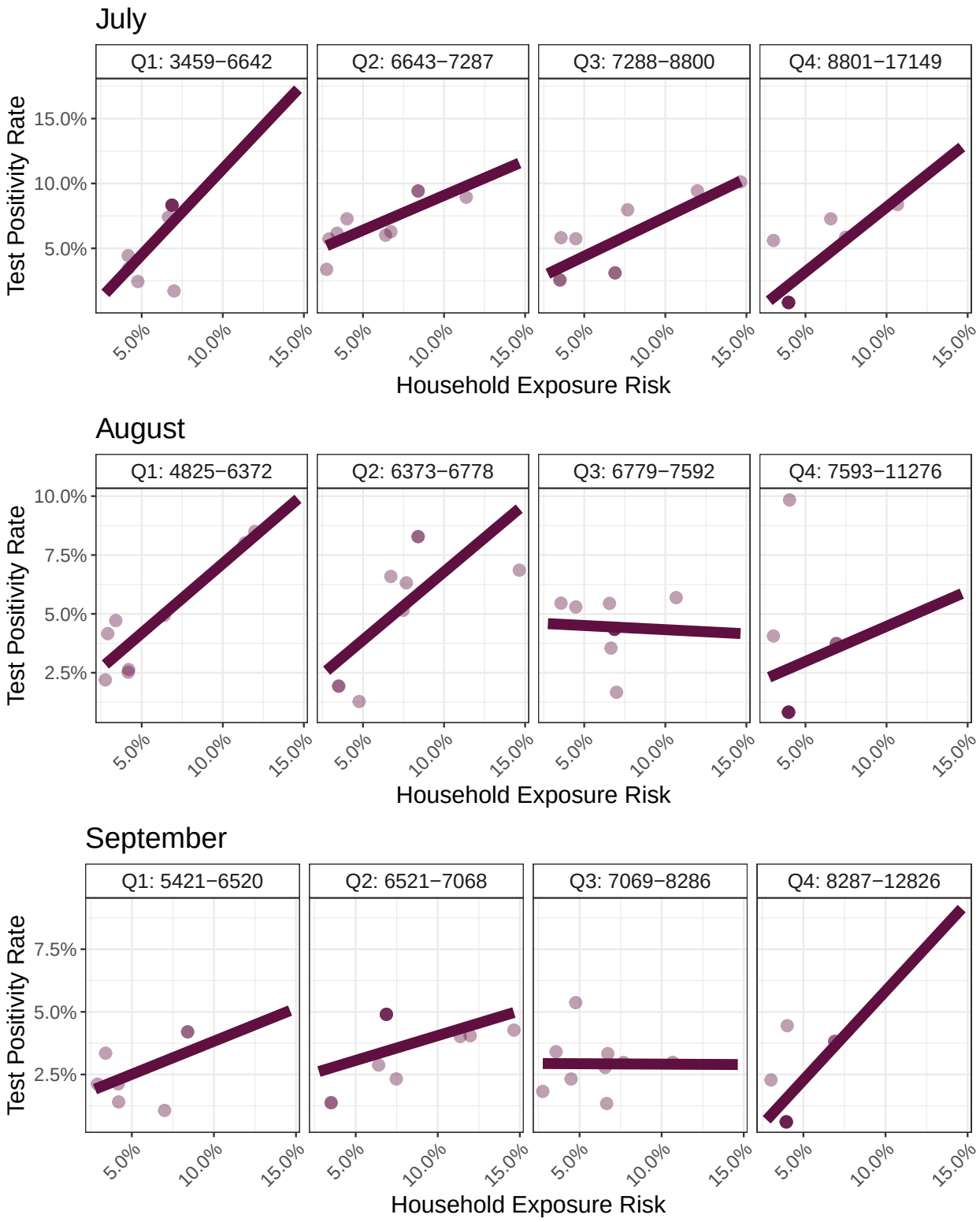
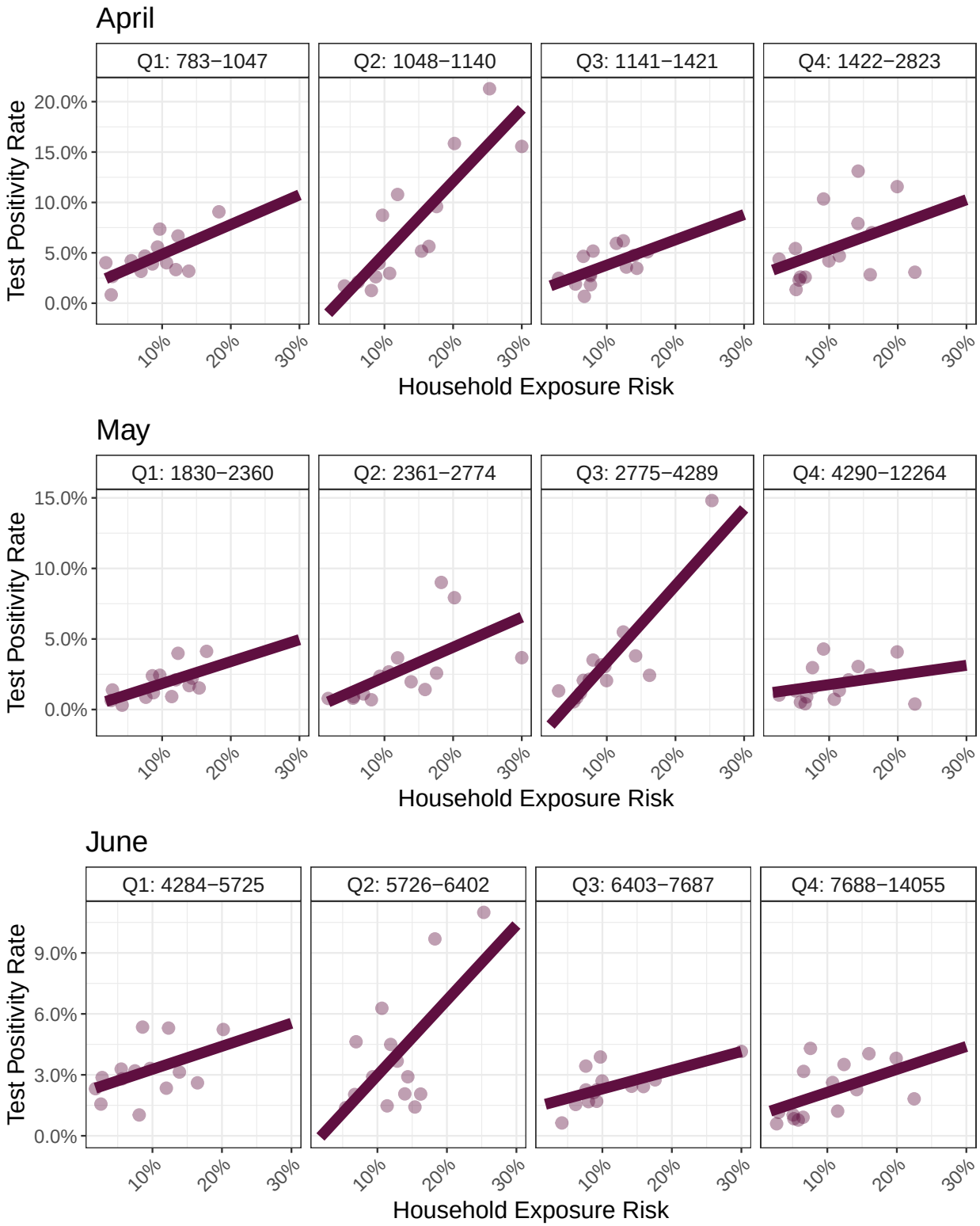

July
Q1: 3459−6642
Q2: 6643−7287
Q3: 7288−8800
Q4: 8801−17149
Test Positivity Rate
15.0%
10.0%
5.0%
Household Exposure Risk
August
Q1: 4825−6372
Q2: 6373−6778
Q3: 6779−7592
Q4: 7593−11276
Test Positivity Rate
10.0%
7.5%
5.0%
2.5%
Household Exposure Risk
September
Q1: 5421−6520
Q2: 6521−7068
Q3: 7069−8286
Q4: 8287−12826
Test Positivity Rate
7.5%
5.0%
2.5%
Household Exposure Risk

Exhibit S7. Household exposure risk vs. test positivity rate, by PUMA and test rate quartile, shown for for April-September for all PUMAs in the Bay Region.



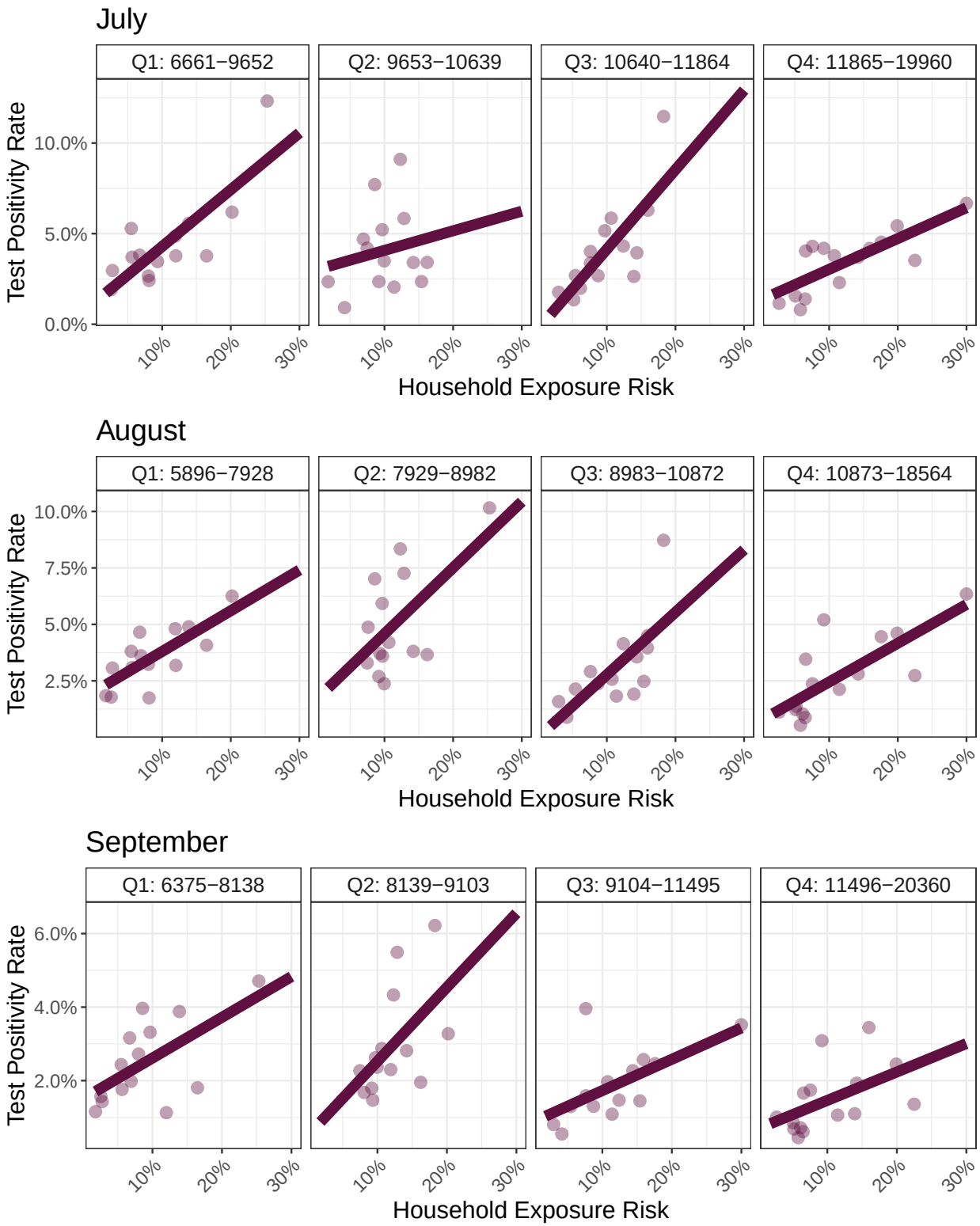
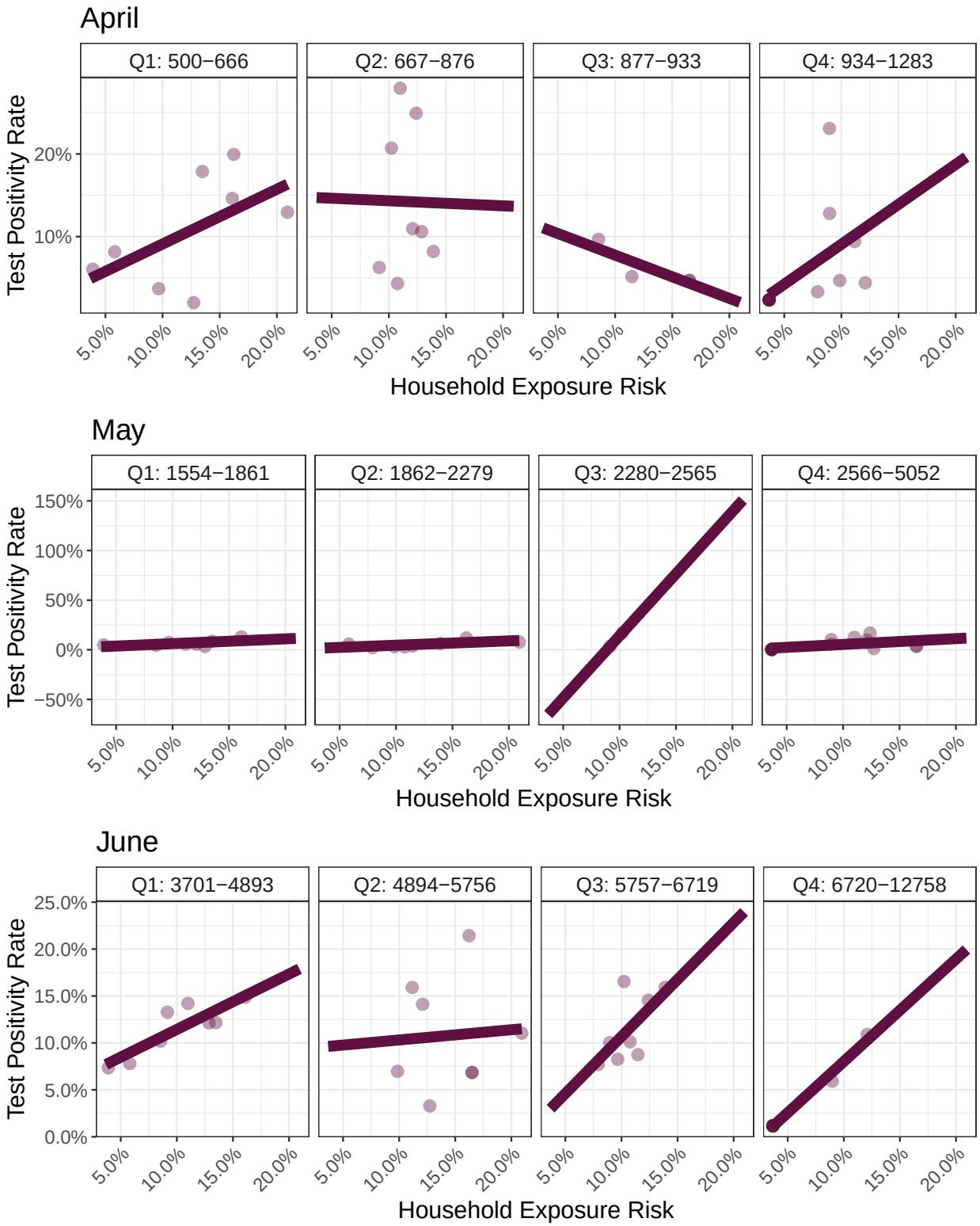

July
Q1: 6661−9652
Q2: 9653−10639
Q3: 10640−11864
Q4: 11865−19960
Test Positivity Rate
10.0%
5.0%
0.0%
10%
20%
30%
Household Exposure Risk
August
Q1: 5896−7928
Q2: 7929−8982
Q3: 8983−10872
Q4: 10873−18564
Test Positivity Rate
10.0%
7.5%
5.0%
2.5%
10%
20%
30%
Household Exposure Risk
September
Q1: 6375−8138
Q2: 8139−9103
Q3: 9104−11495
Q4: 11496−20360
Test Positivity Rate
6.0%
4.0%
2.0%
10%
20%
30%
Household Exposure Risk

Exhibit S8. Household exposure risk vs. test positivity rate, by PUMA and test rate quartile, shown for for April-September for all PUMAs in the Central Region.



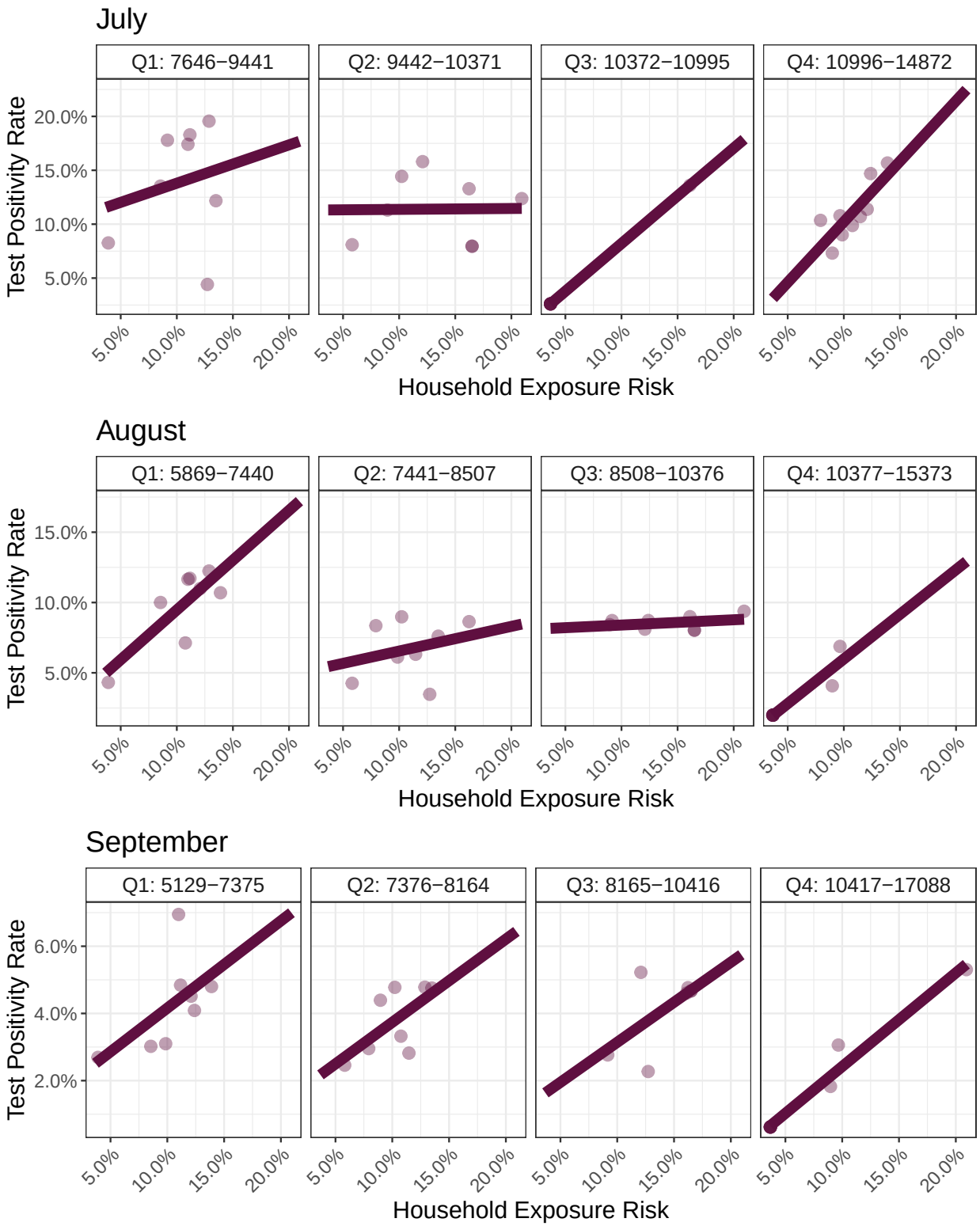
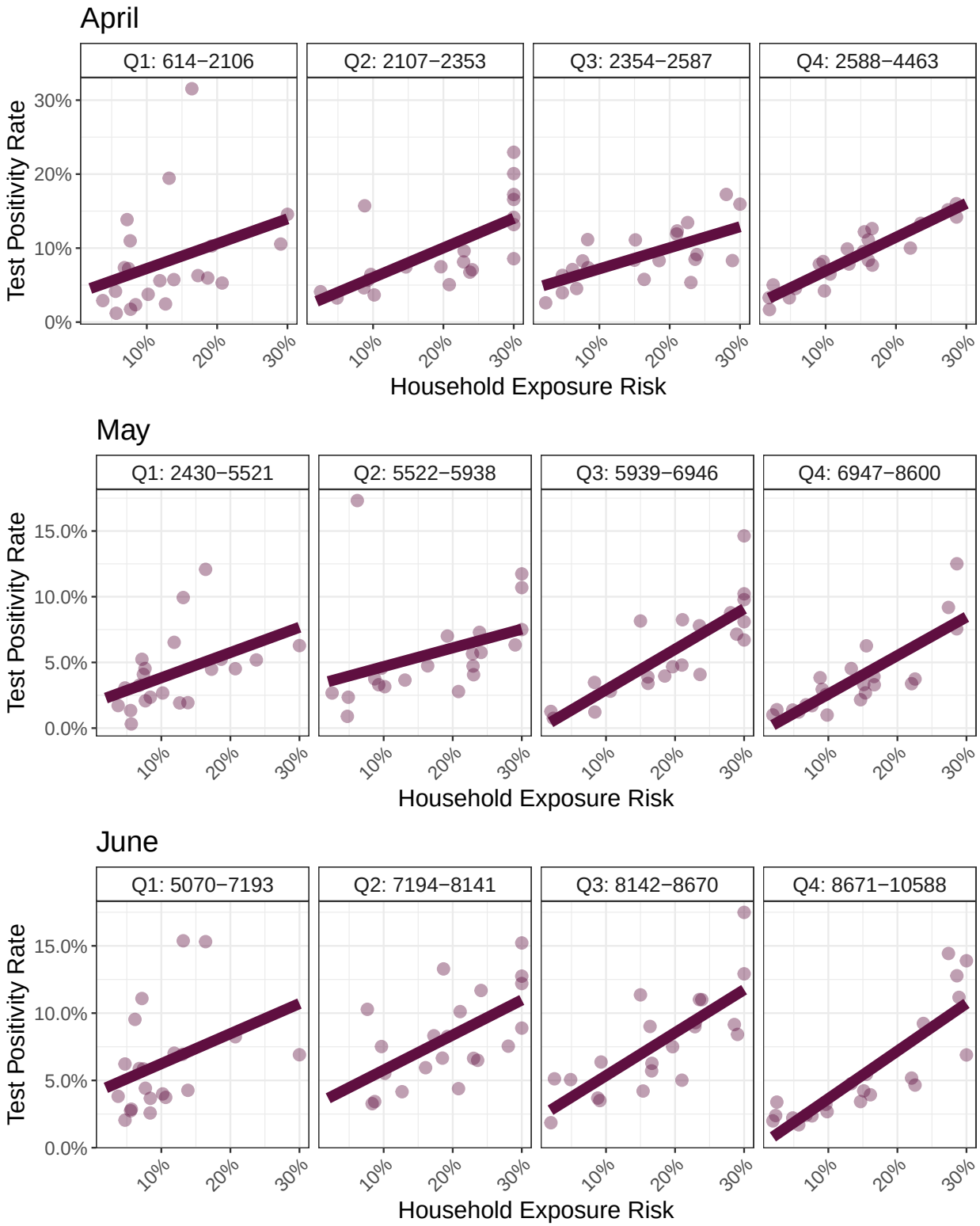

July
Q1: 7646−9441
Q2: 9442−10371
Q3: 10372−10995
Q4: 10996−14872
Test Positivity Rate
20.0%
15.0%
10.0%
5.0%
Household Exposure Risk
August
Q1: 5869−7440
Q2: 7441−8507
Q3: 8508−10376
Q4: 10377−15373
Test Positivity Rate
15.0%
10.0%
5.0%
Household Exposure Risk
September
Q1: 5129−7375
Q2: 7376−8164
Q3: 8165−10416
Q4: 10417−17088
Test Positivity Rate
6.0%
4.0%
2.0%
Household Exposure Risk

Exhibit S9. Household exposure risk vs. test positivity rate, by PUMA and test rate quartile, shown for for April-September for all PUMAs in the Upper Southern Region.



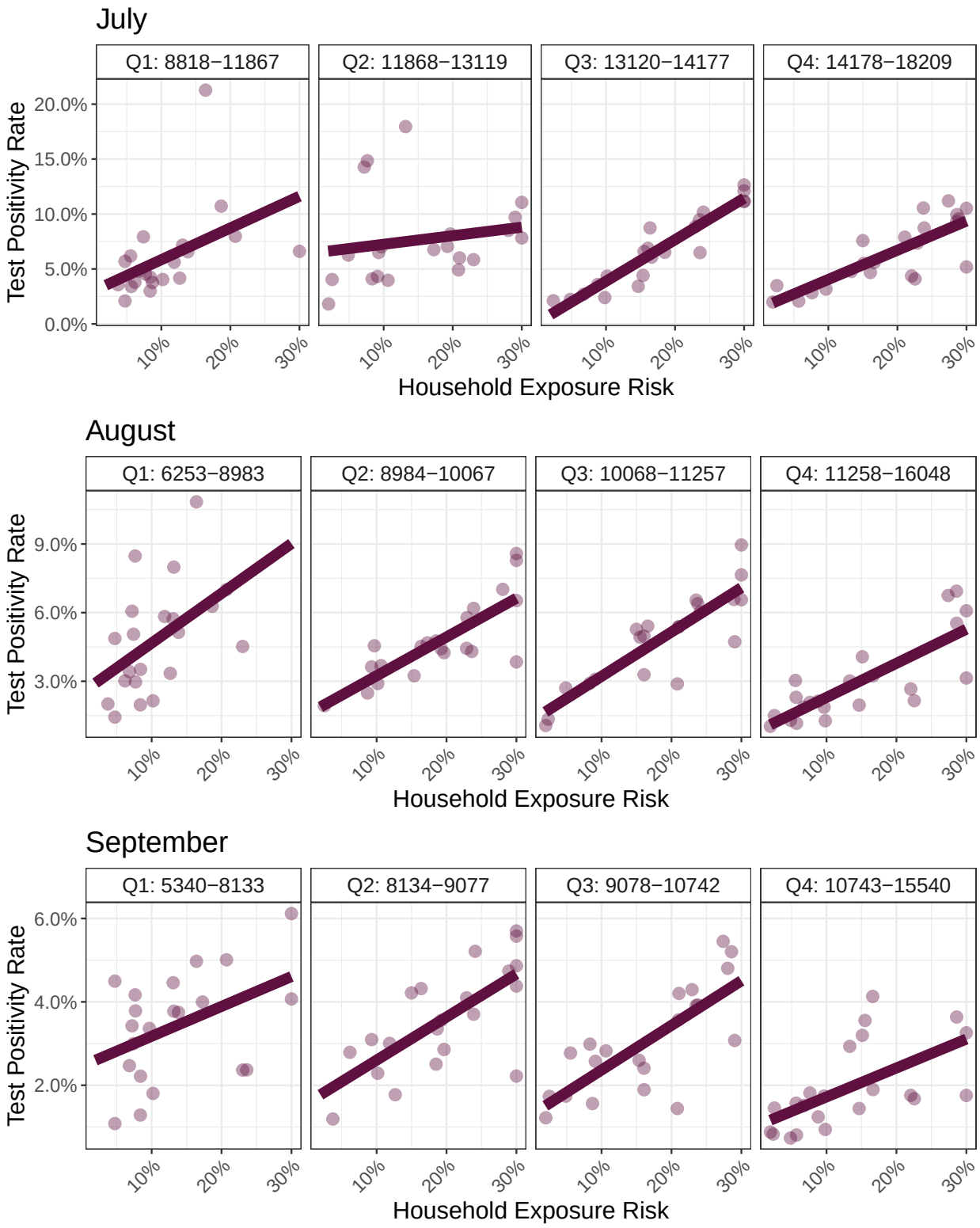
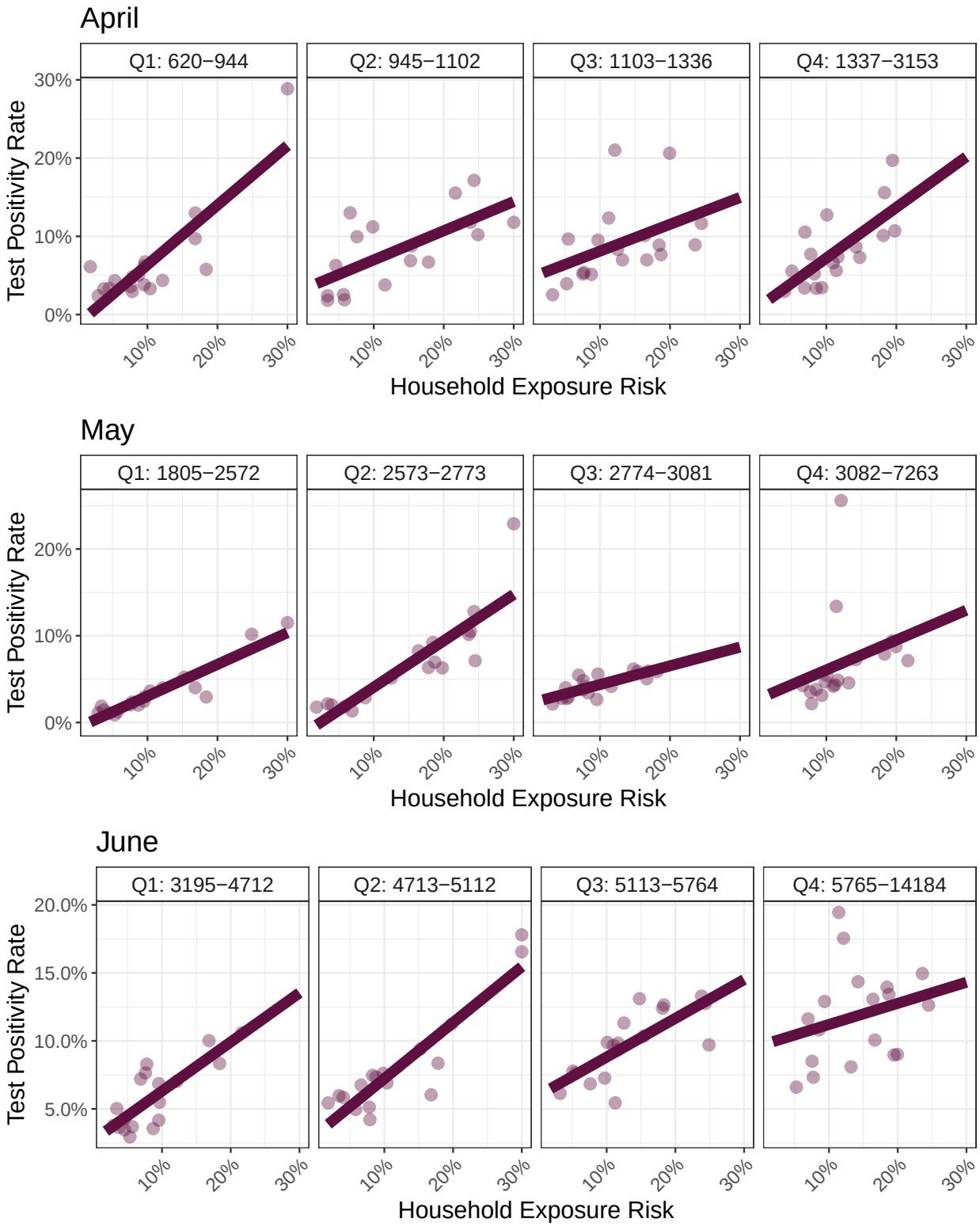

July
Q1: 8818−11867
Q2: 11868−13119
Q3: 13120−14177
Q4: 14178−18209
Test Positivity Rate
20.0%
15.0%
10.0%
5.0%
0.0%
10%
20%
30%
Household Exposure Risk
August
Q1: 6253−8983
Q2: 8984−10067
Q3: 10068−11257
Q4: 11258−16048
Test Positivity Rate
9.0%
6.0%
3.0%
Household Exposure Risk
September
Q1: 5340−8133
Q2: 8134−9077
Q3: 9078−10742
Q4: 10743−15540
Test Positivity Rate
6.0%
4.0%
2.0%
Household Exposure Risk

Exhibit S10. Household exposure risk vs. test positivity rate, by PUMA and test rate quartile, shown for for April-September for all PUMAs in the Lower Southern Region.



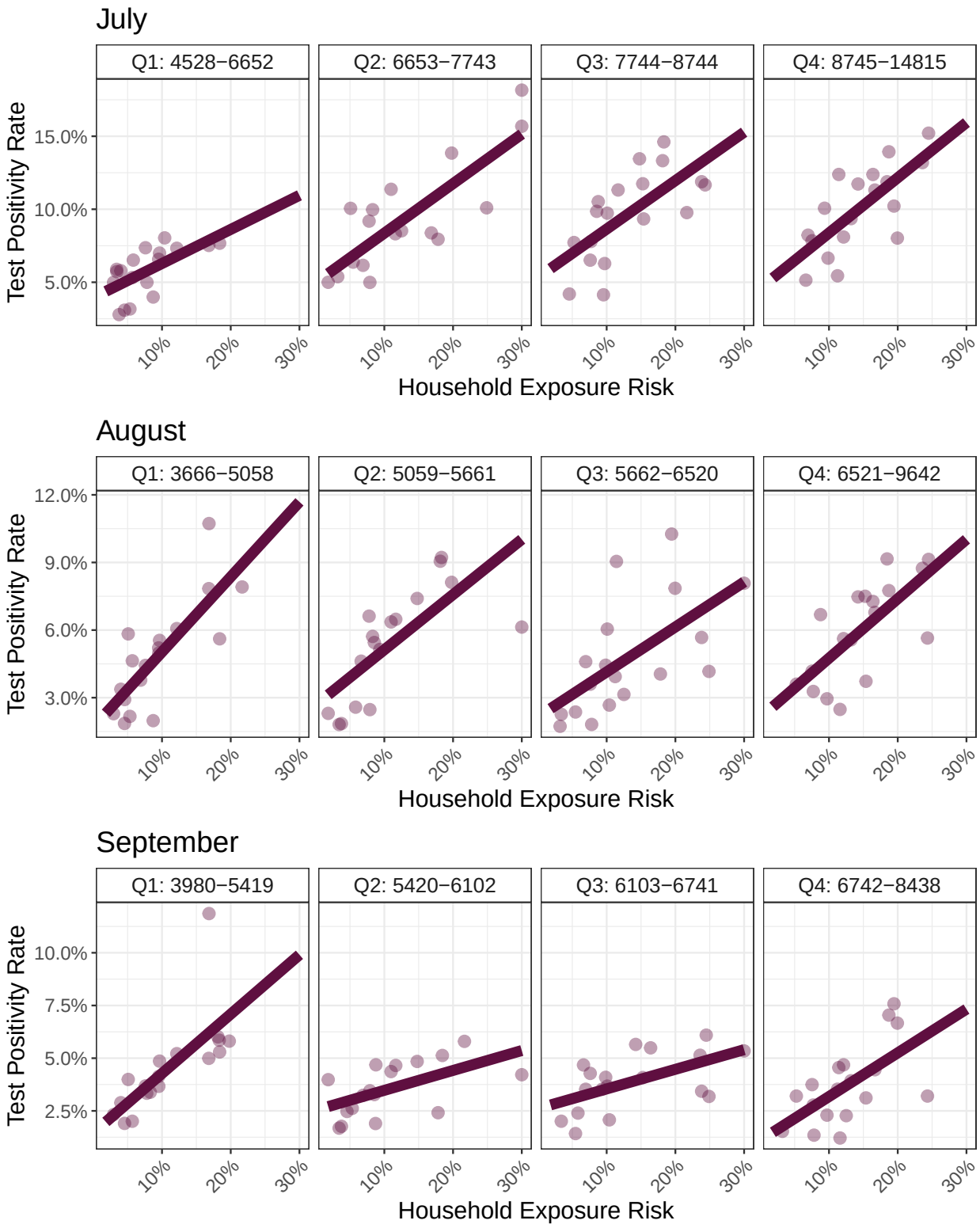

July
Q1: 4528−6652
Q2: 6653−7743
Q3: 7744−8744
Q4: 8745−14815
Test Positivity Rate
15.0%
10.0%
5.0%
10%
20%
30%
Household Exposure Risk
August
Q1: 3666−5058
Q2: 5059−5661
Q3: 5662−6520
Q4: 6521−9642
Test Positivity Rate
12.0%
9.0%
6.0%
3.0%
10%
20%
30%
Household Exposure Risk
September
Q1: 3980−5419
Q2: 5420−6102
Q3: 6103−6741
Q4: 6742−8438
Test Positivity Rate
10.0%
7.5%
5.0%
2.5%
10%
20%
30%
Household Exposure Risk

Exhibit S11. Test positivity rate, by race and colored by percent of people living in household with high exposure risk, shown for April through September and grouped by region.

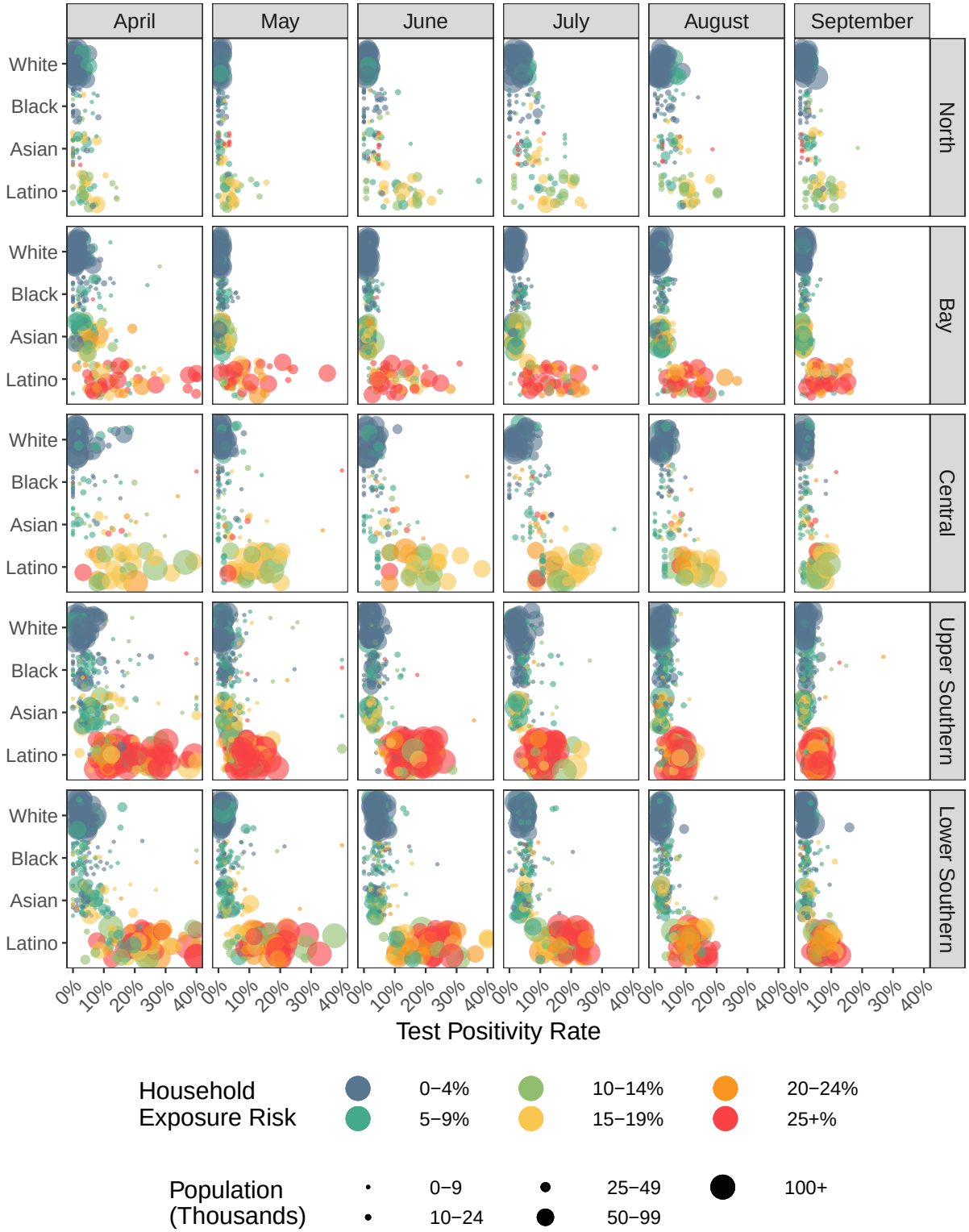


Exhibit S12. Test positivity rate, by race and colored by test rate per 100,000 population, shown for April through September and grouped by region.

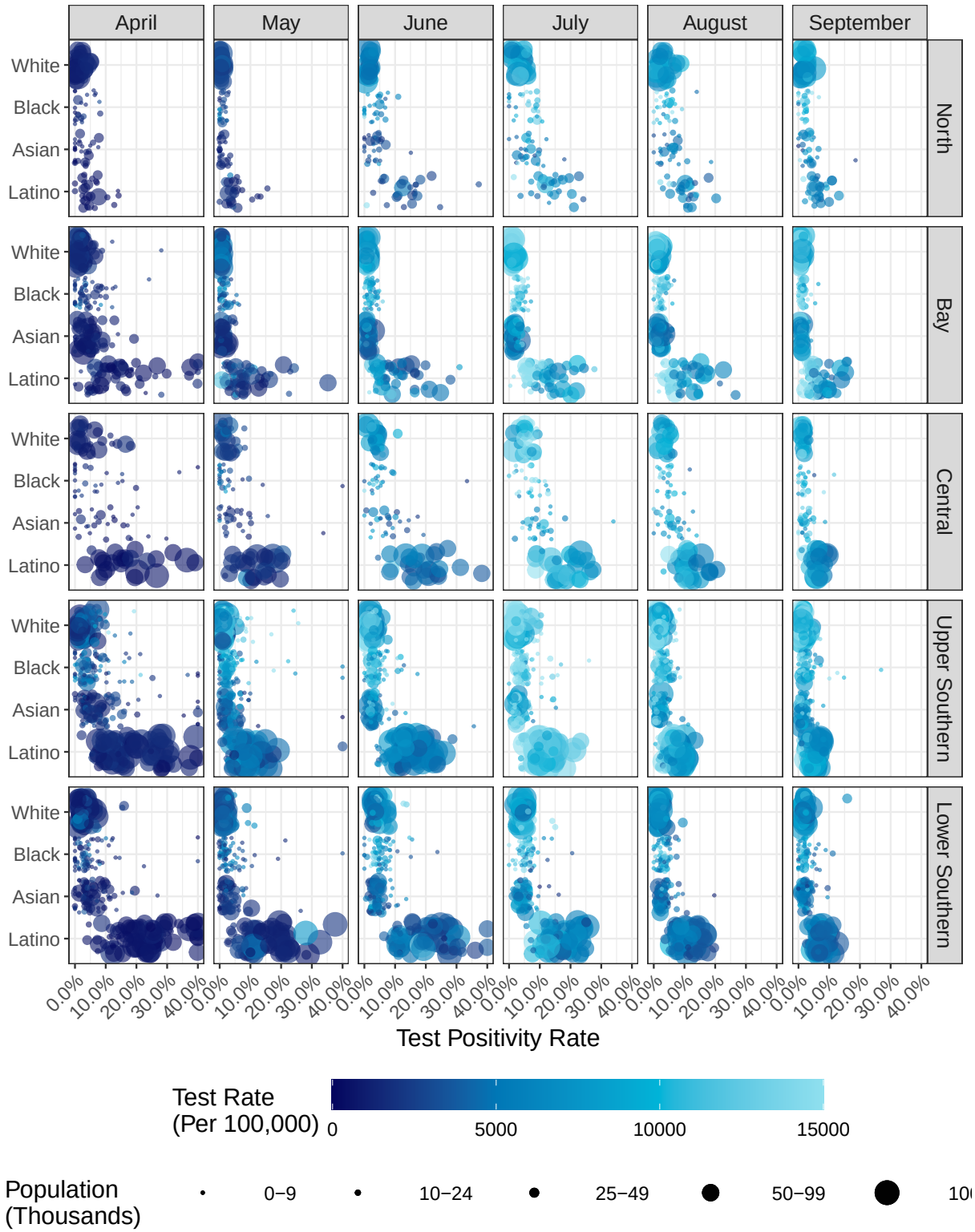


Exhibit S13. Percent of tests with PUMA-level residential geolocation available, out of all tests in a county.

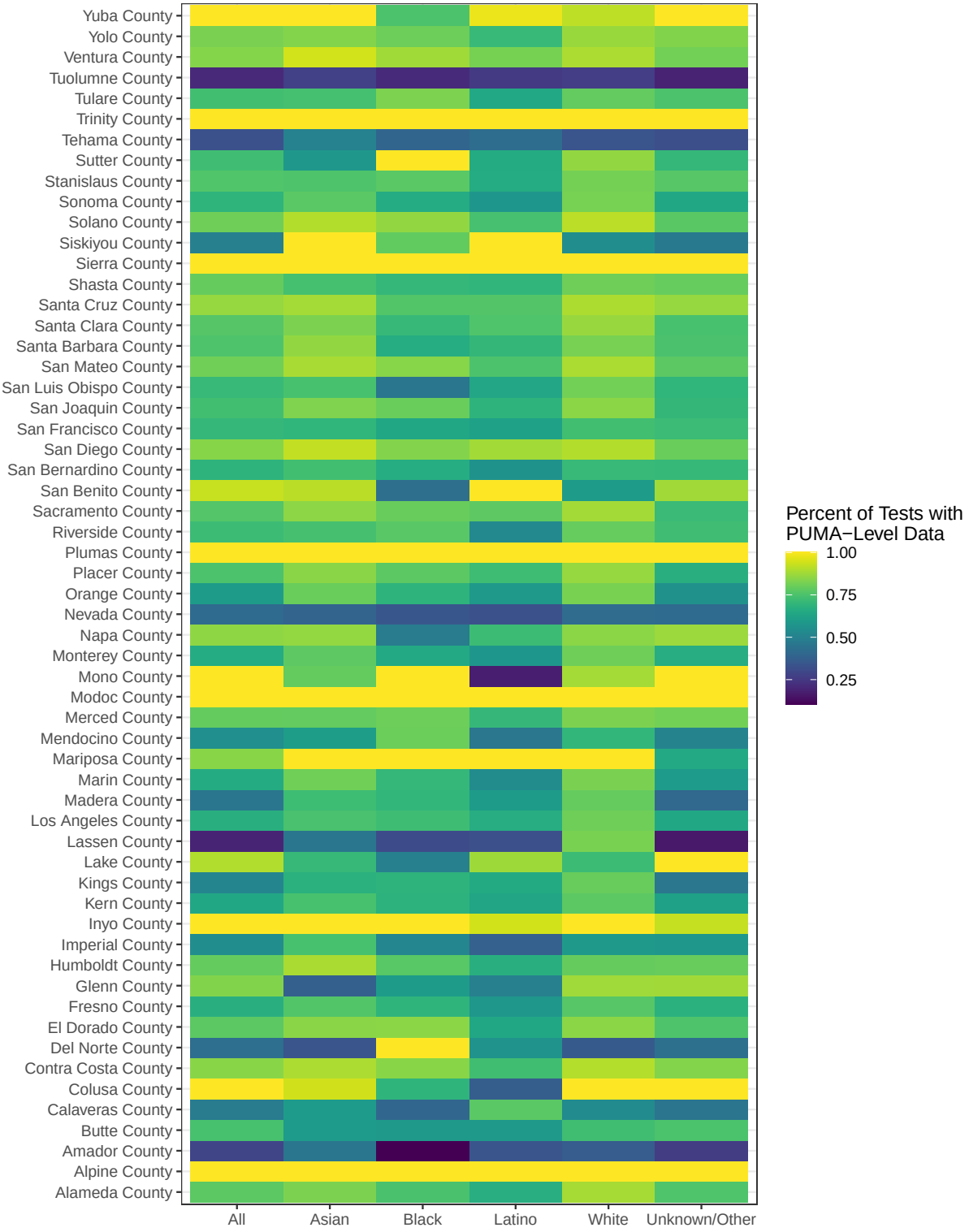


Exhibit S14. Percent of cases with PUMA-level residential geolocation available, out of all cases in a county.

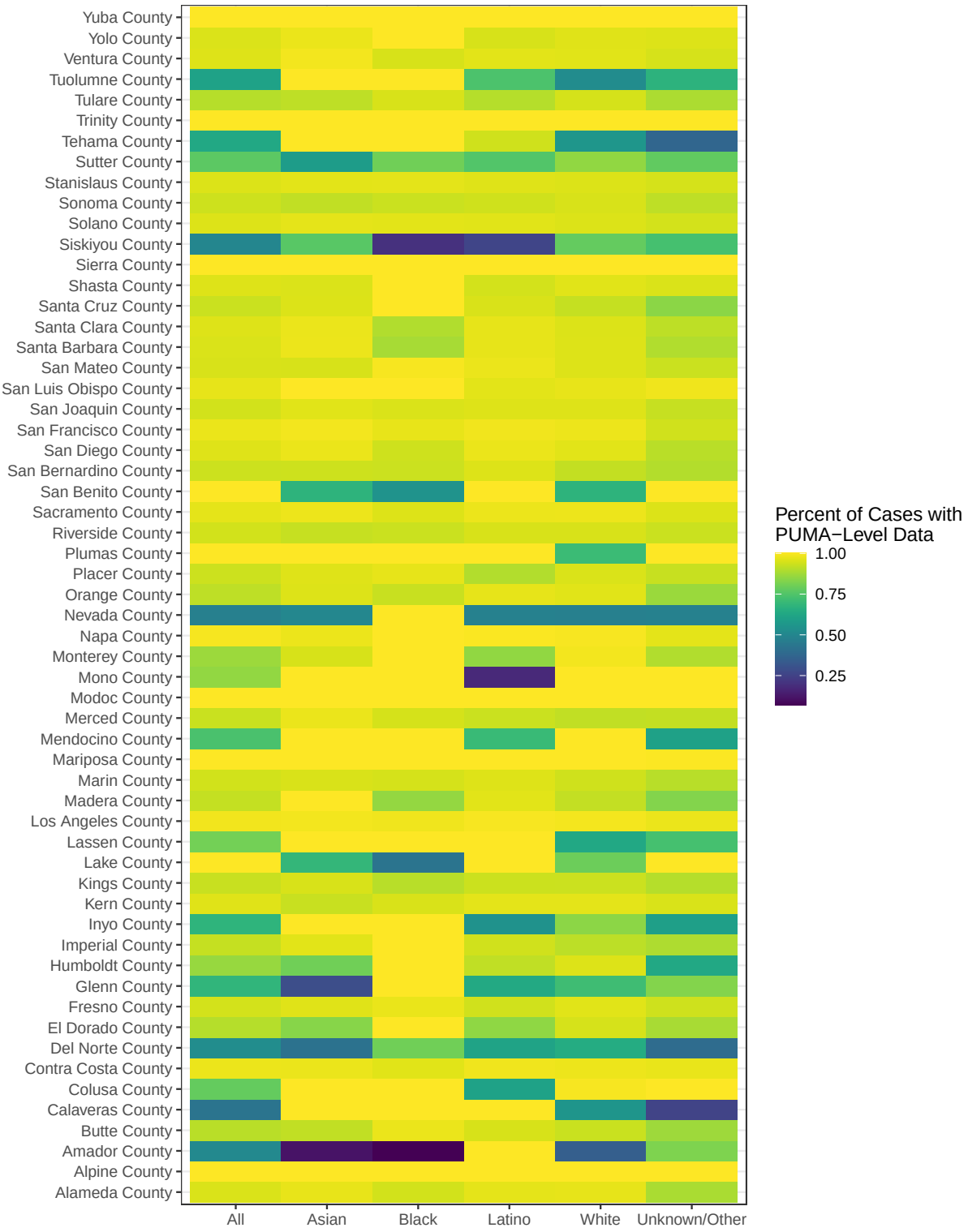


Exhibit S15. Manuscript Exhibit 3 with separate panels for test rate quartiles.

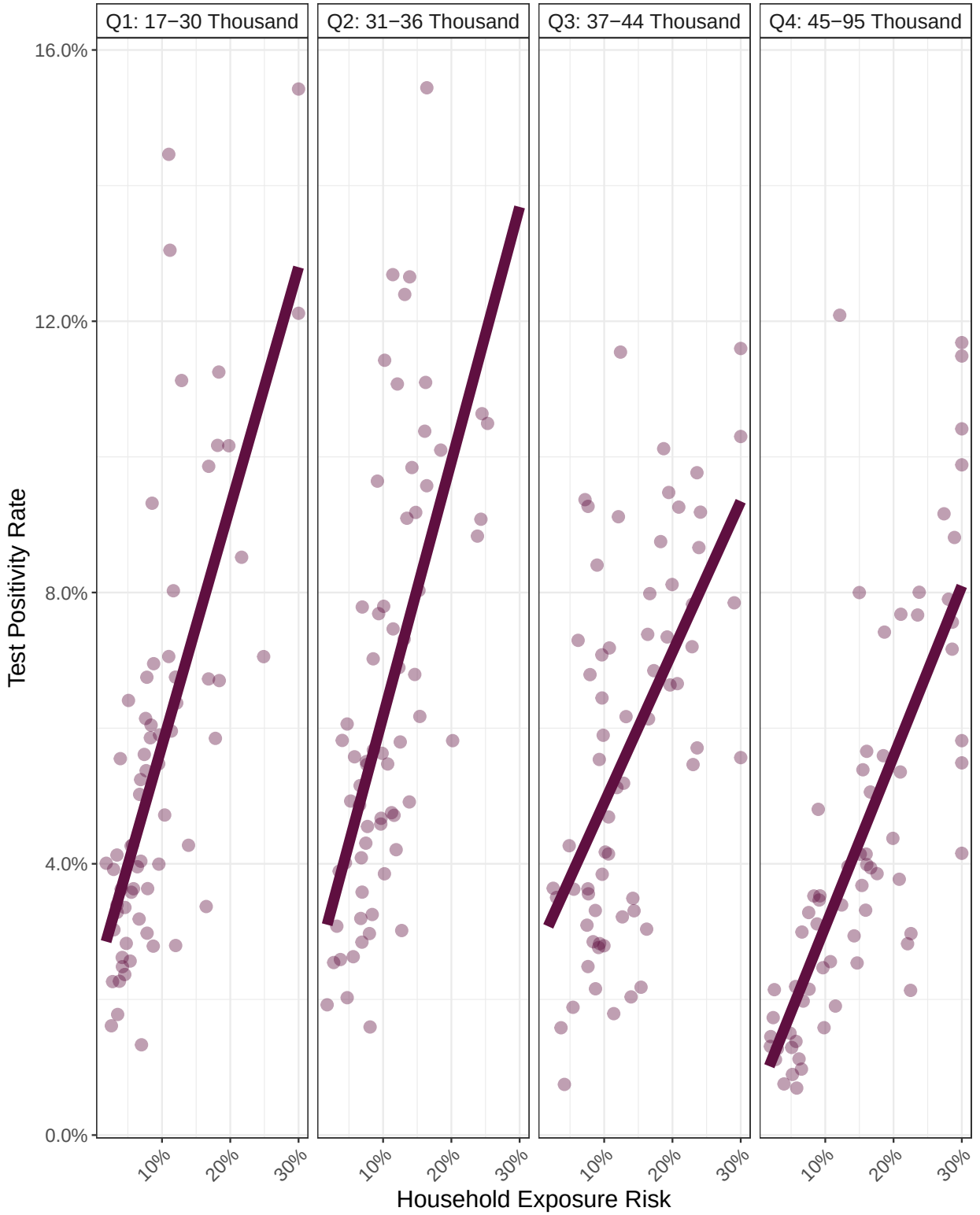


Exhibit S16. Sensitivity analysis for manuscript Exhibit 3 distributing cases and tests with missing data based on population size.

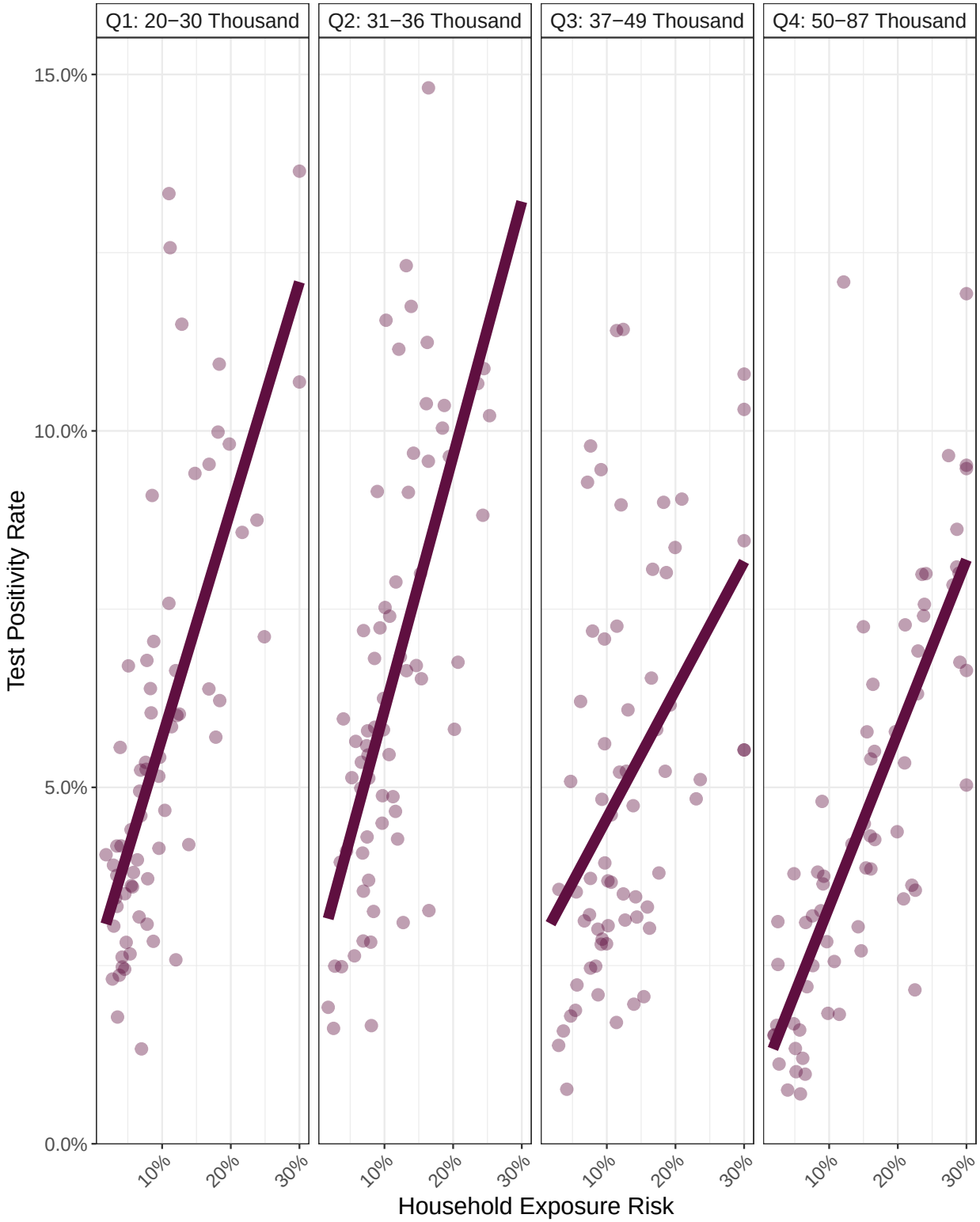


Exhibit S17. Sensitivity analysis for manuscript Exhibit 4 distributing cases and tests with missing data based on population size.

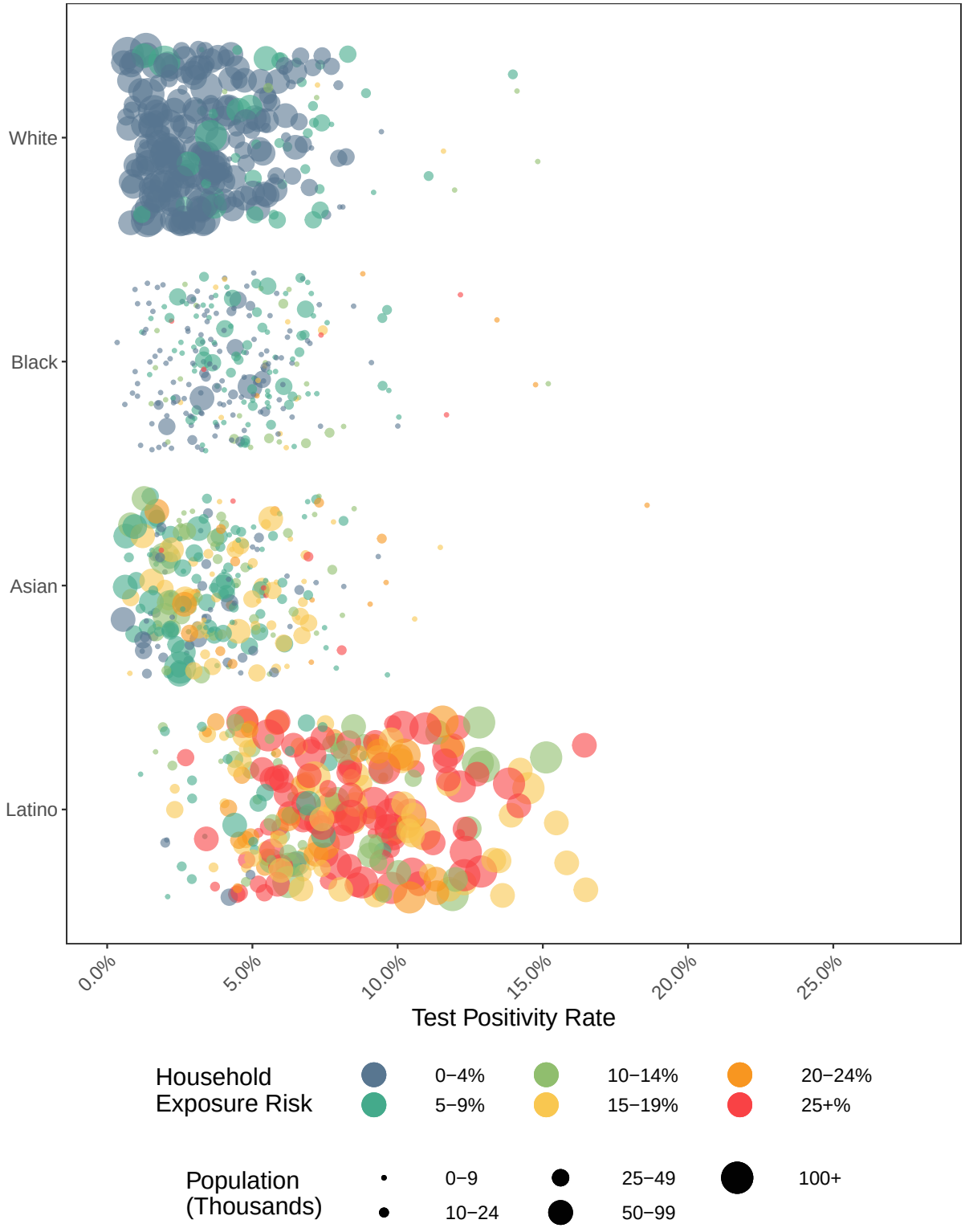


Exhibit S18. Sensitivity analysis for manuscript Exhibit 3 using only data with complete geolocation information.

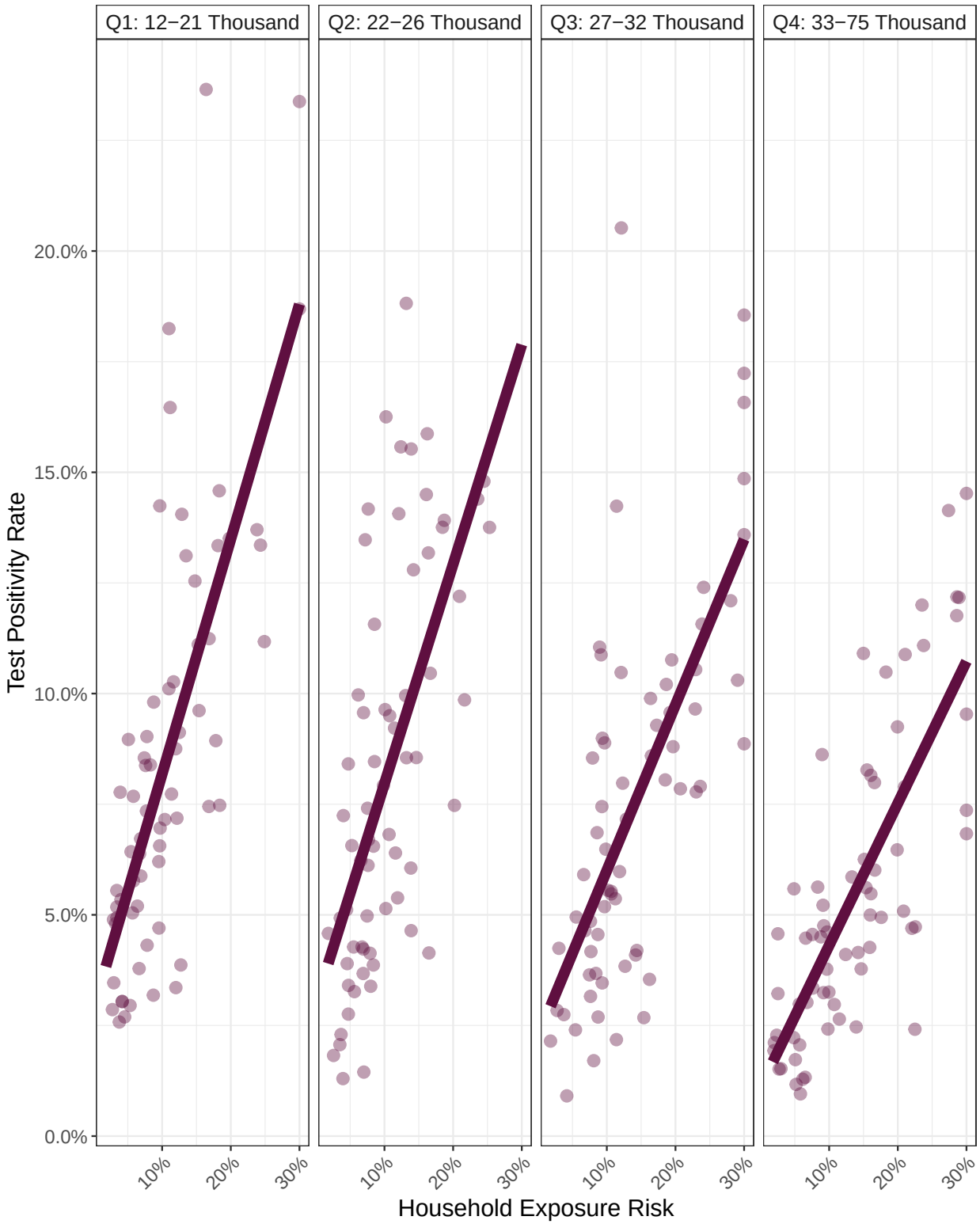


Exhibit S19. Sensitivity analysis for manuscript Exhibit 4 using only data with complete geolocation information and known race/ethnicity.

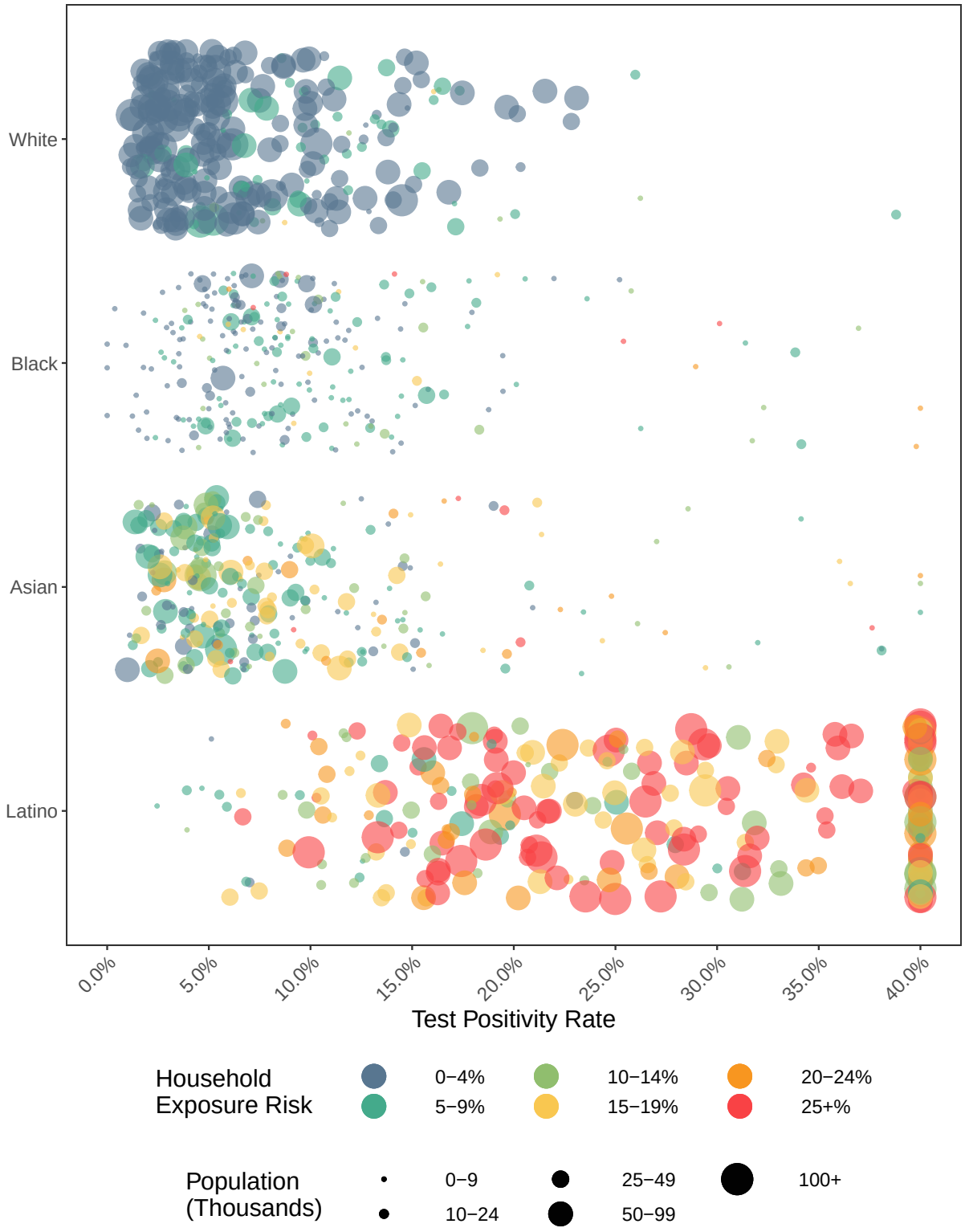


Exhibit S20. Household exposure risk vs. test rate (per 100,000 population), by PUMA, shown for April through September and grouped by region.

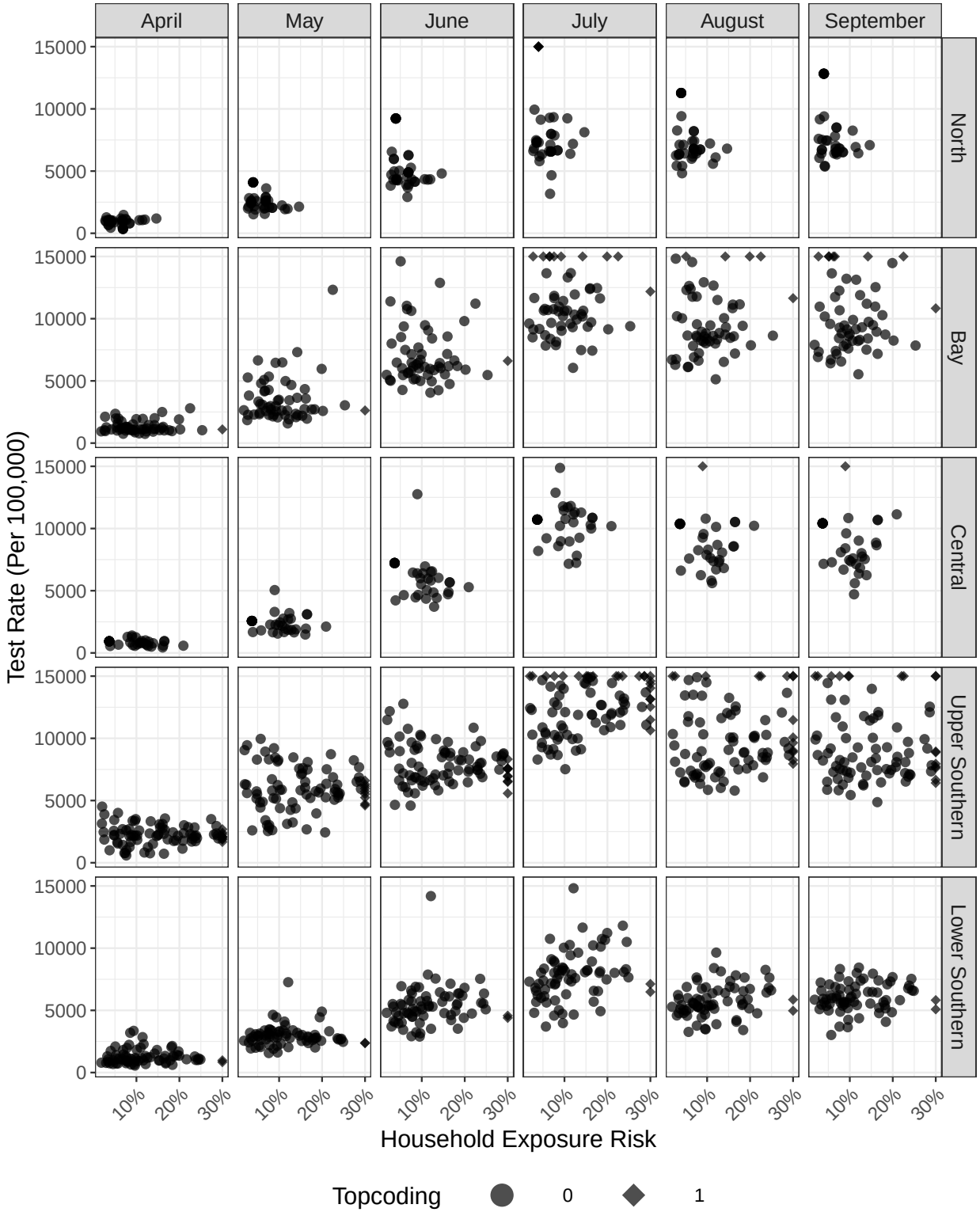
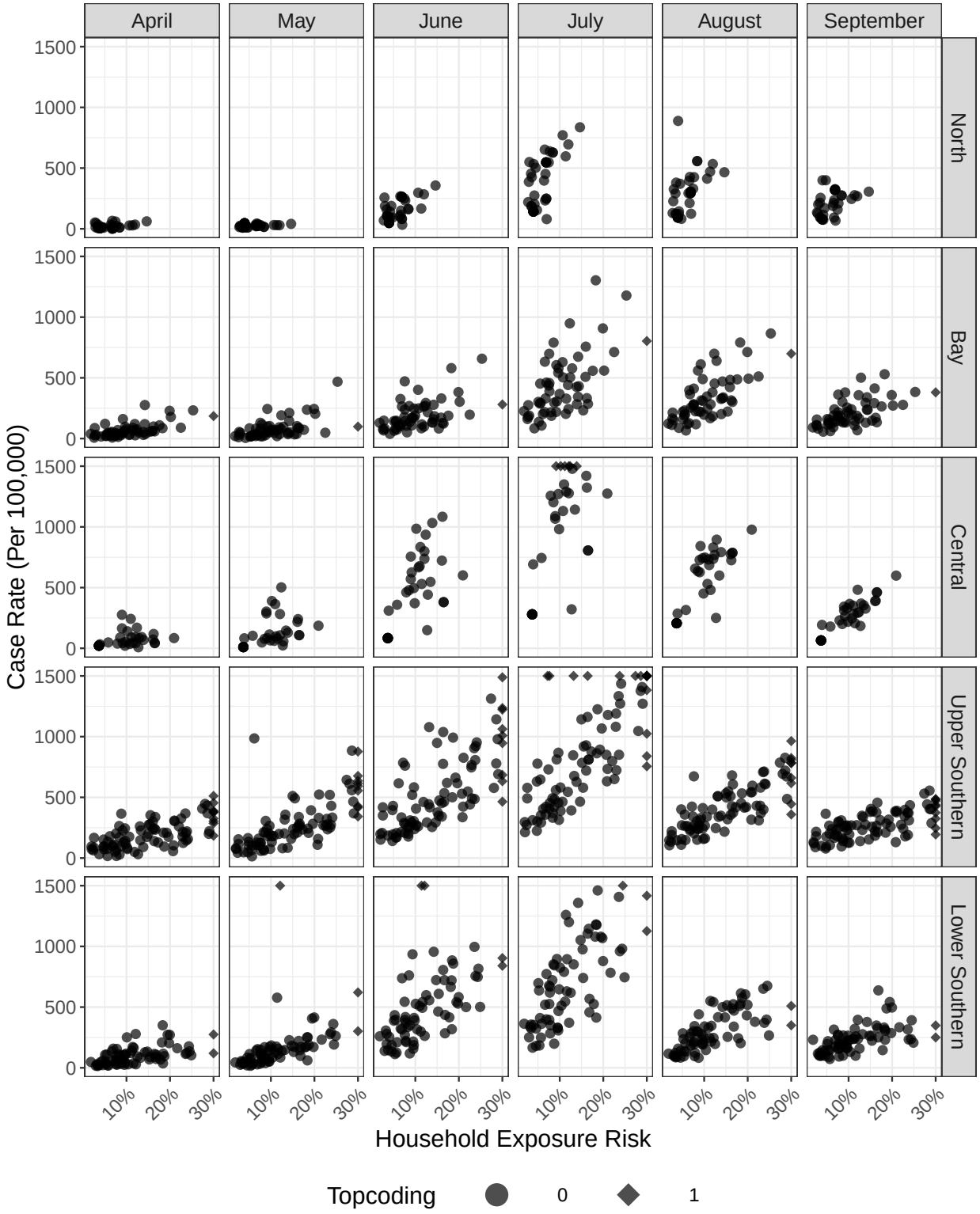
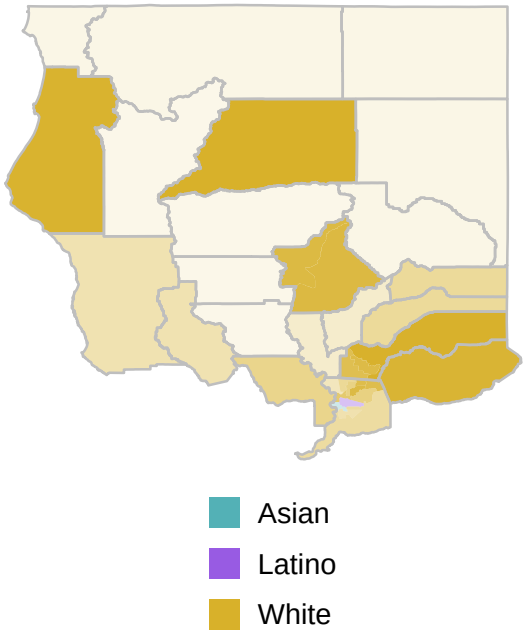


Exhibit S21. Household exposure risk vs. case rate (per 100,000 population), by PUMA, shown for April through September and grouped by region.

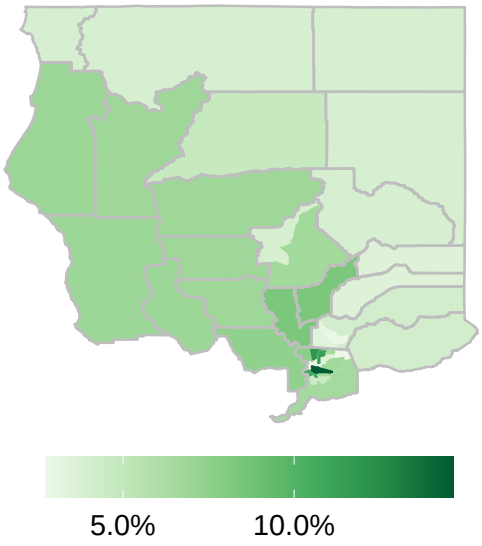


North

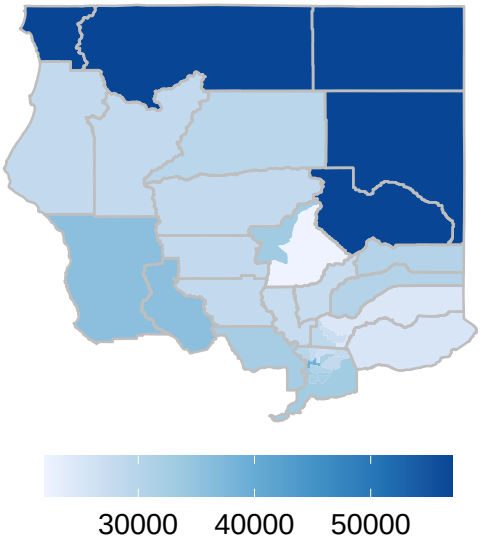
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



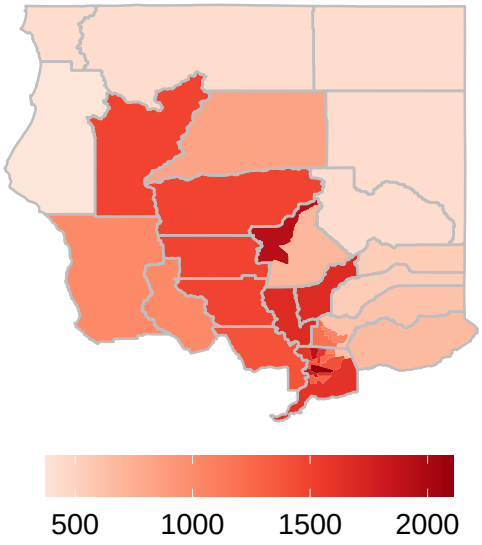
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

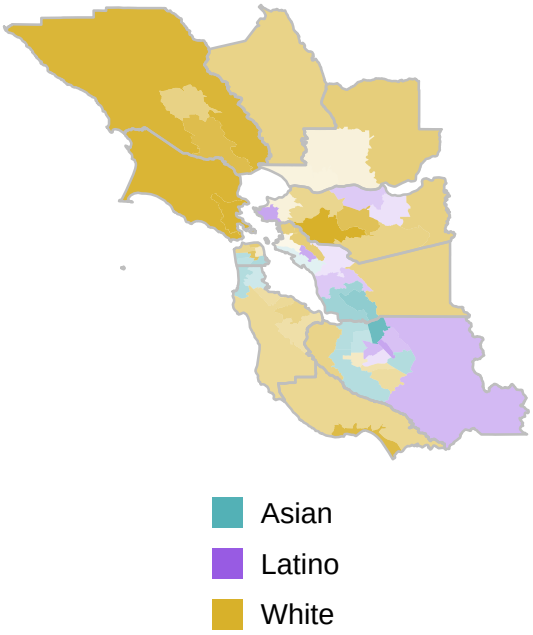


Case Rate
Per 100,000 Population

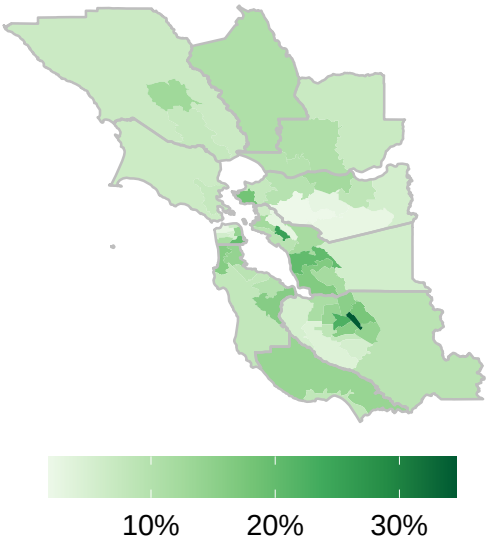


Bay

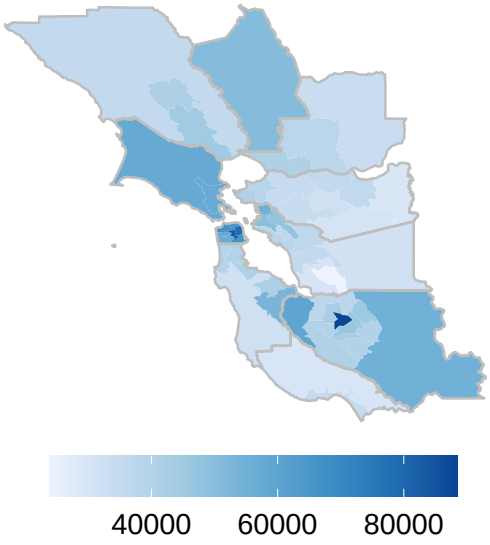
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



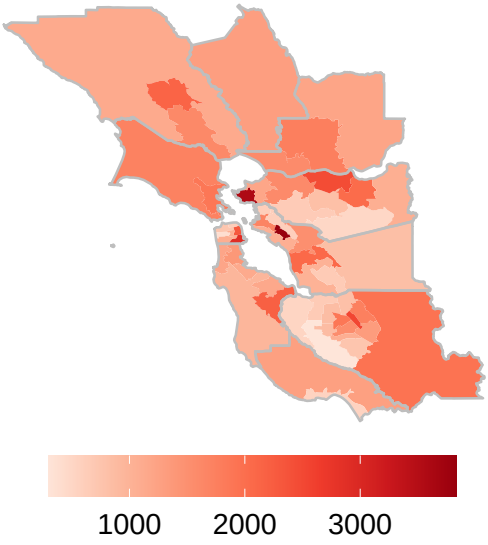
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

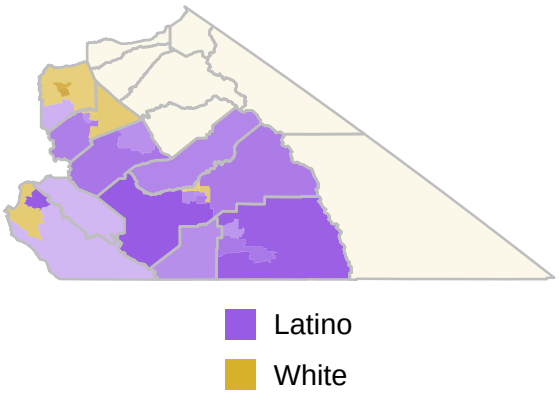


Case Rate
Per 100,000 Population

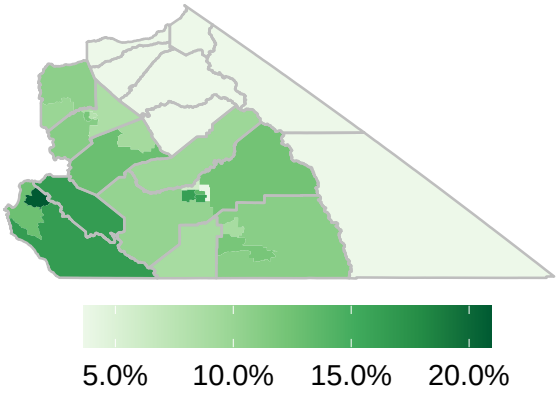


Central

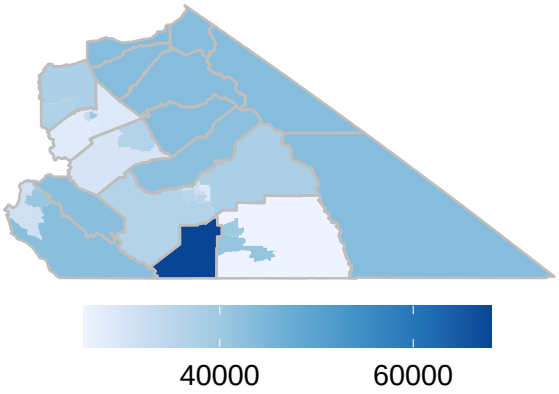
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



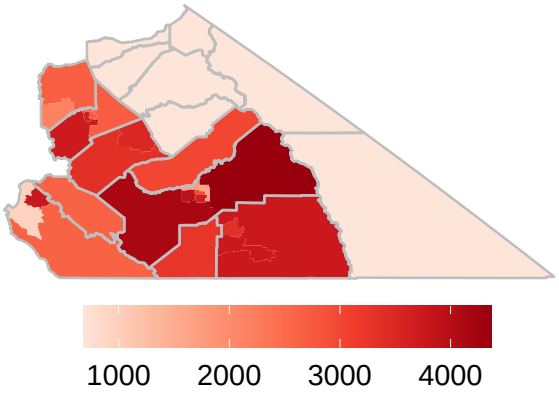
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

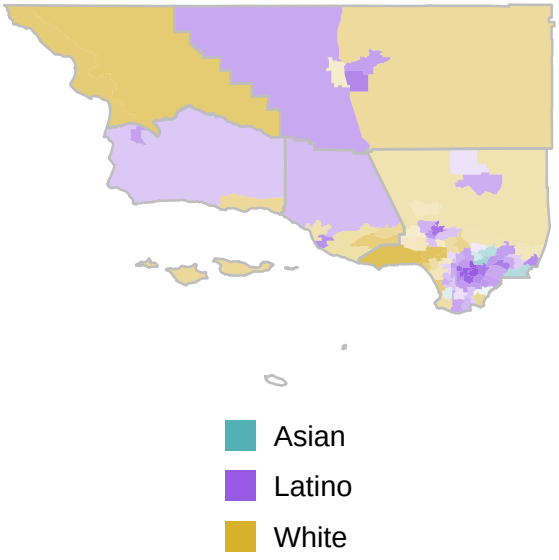


Case Rate
Per 100,000 Population

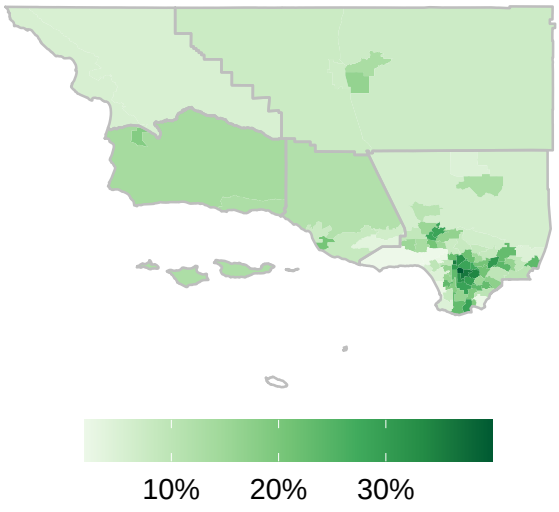


Upper Southern

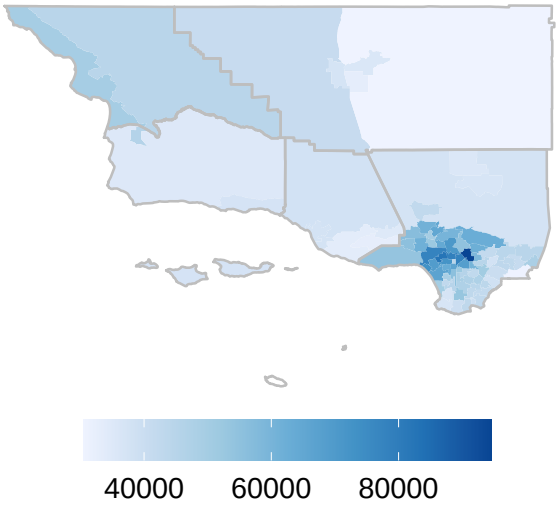
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



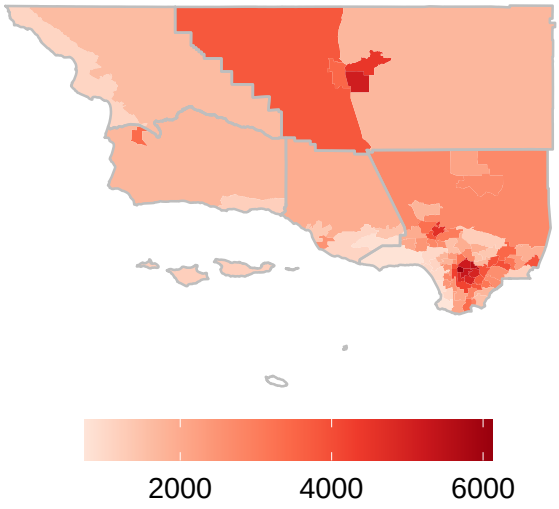
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

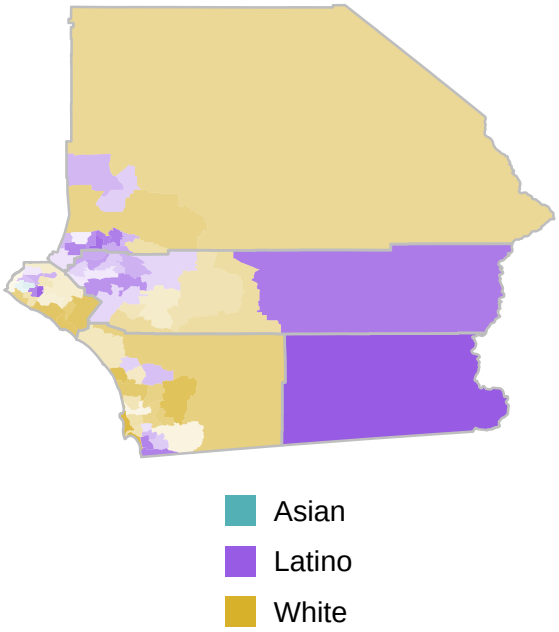


Case Rate
Per 100,000 Population

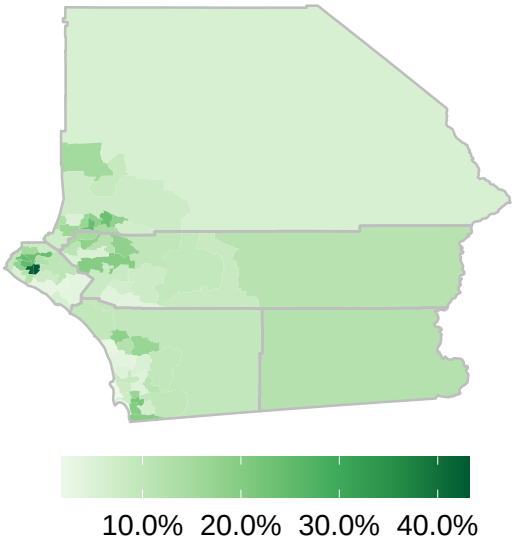


Lower Southern

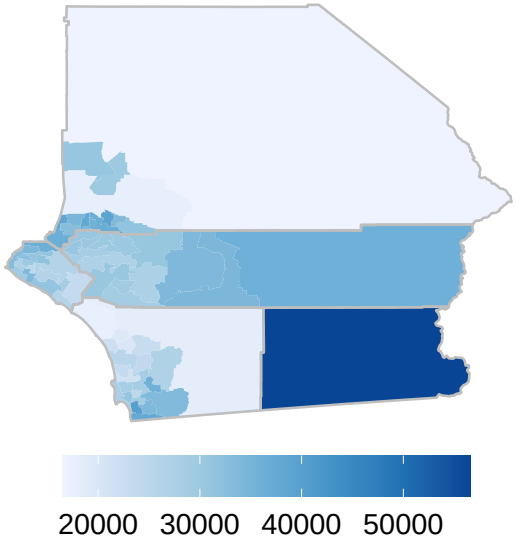
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



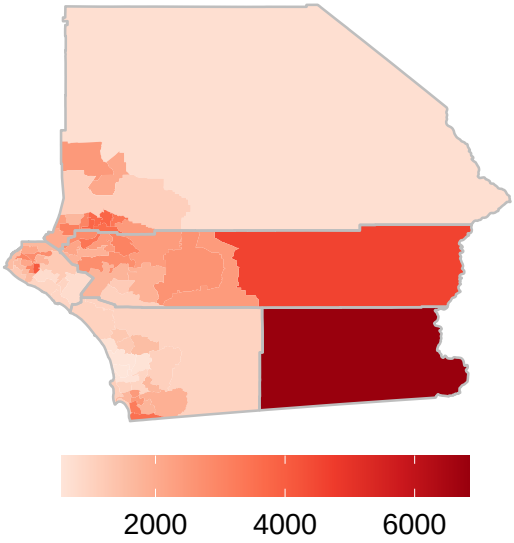
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

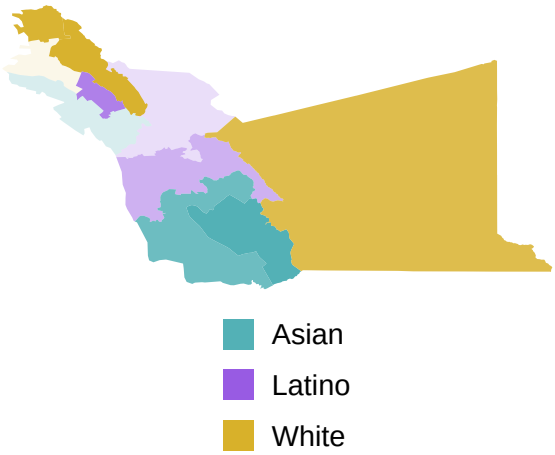


Case Rate
Per 100,000 Population

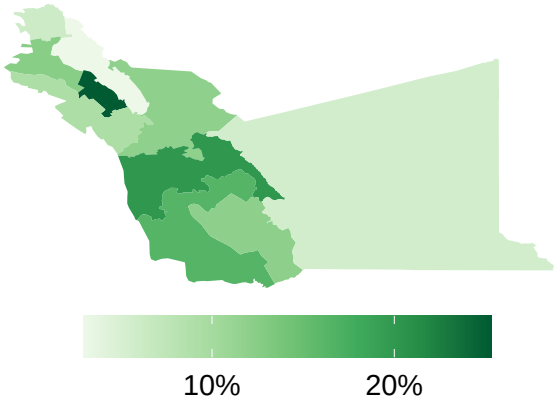


Alameda County

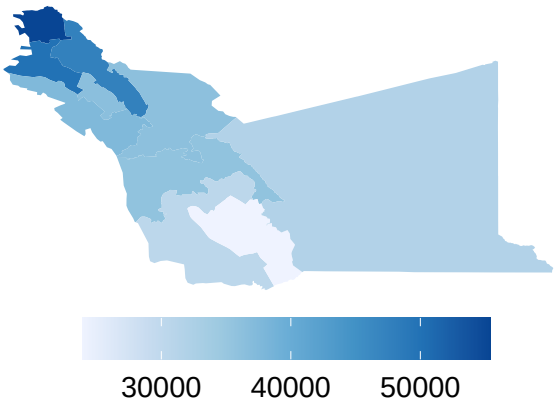
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



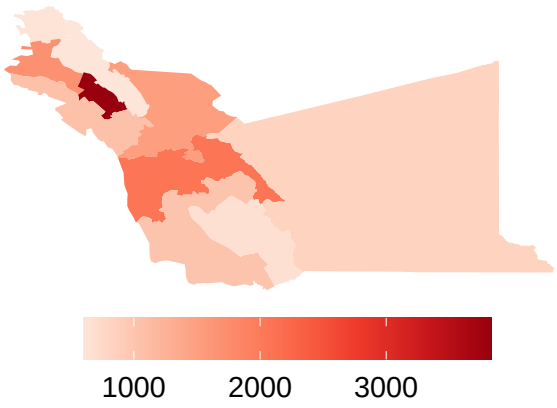
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

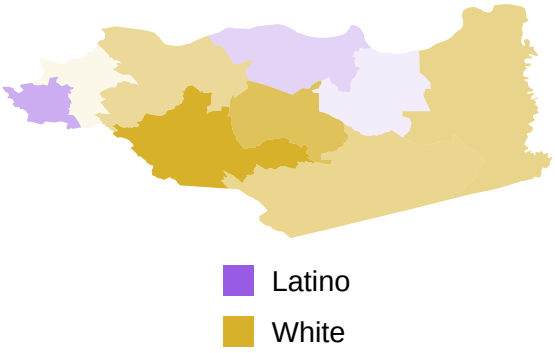


Case Rate
Per 100,000 Population

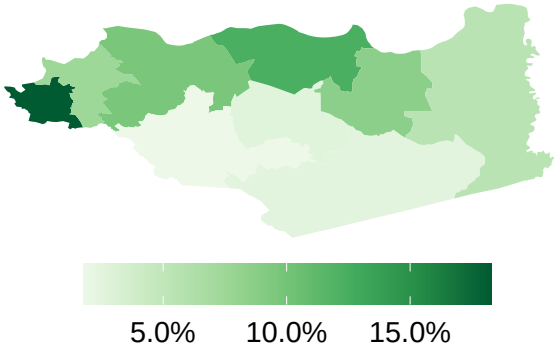


Contra Costa County

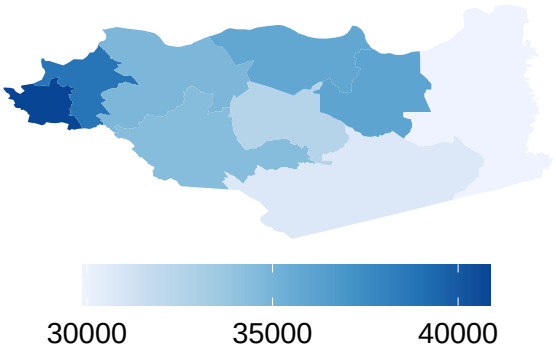
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



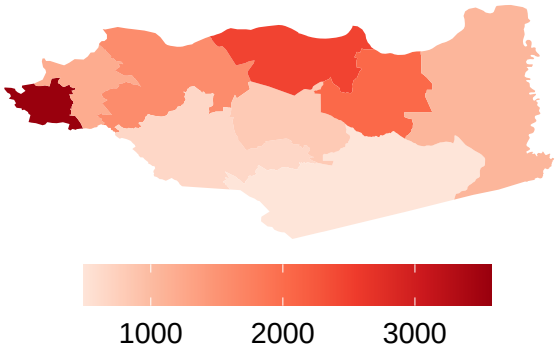
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

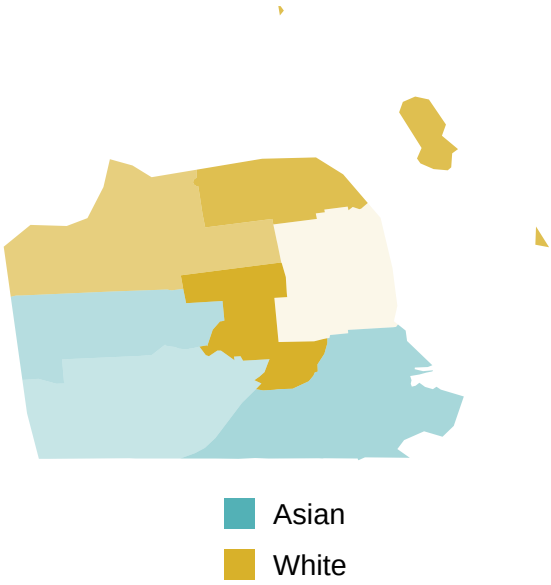


Case Rate
Per 100,000 Population

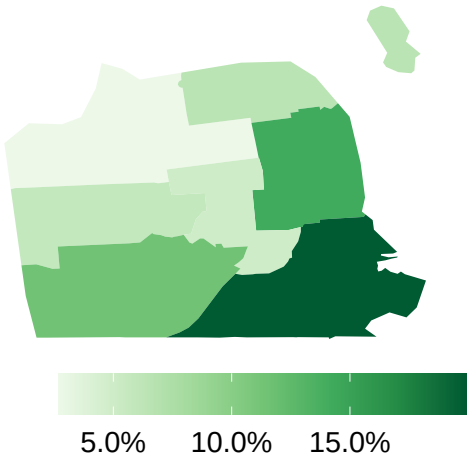


San Francisco County

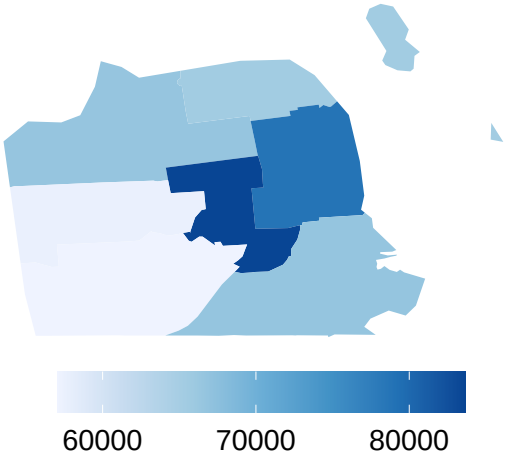
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



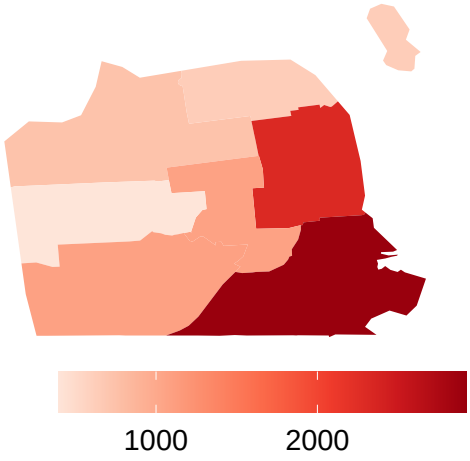
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

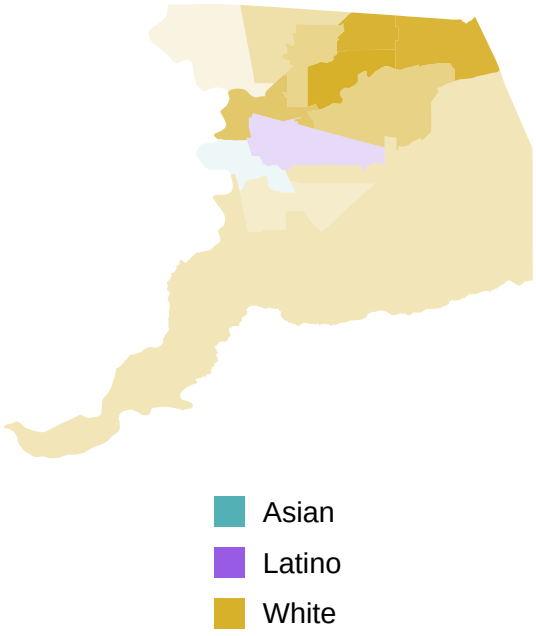


Case Rate
Per 100,000 Population

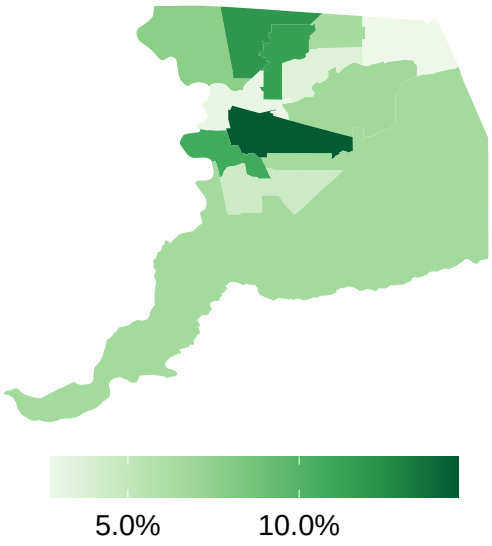


Sacramento County

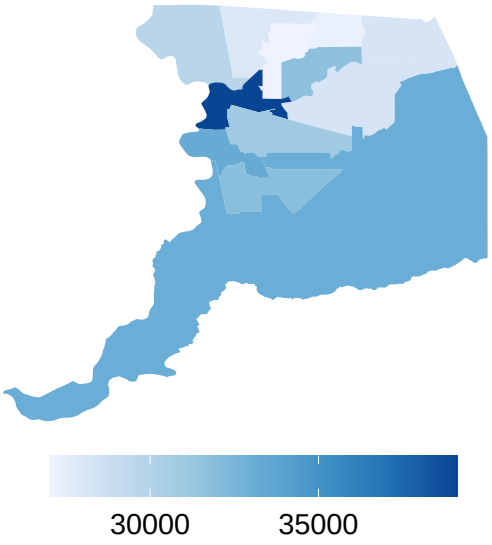
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



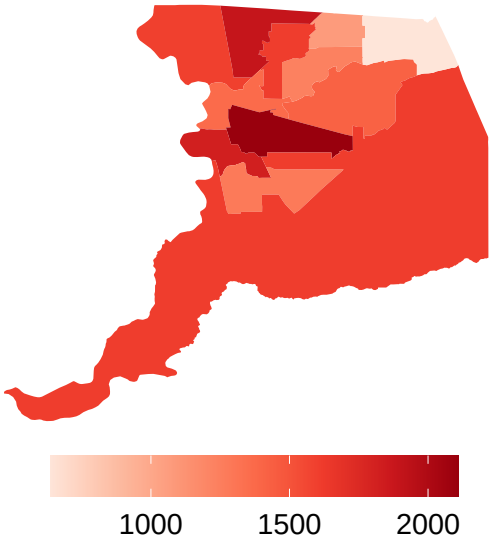
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population



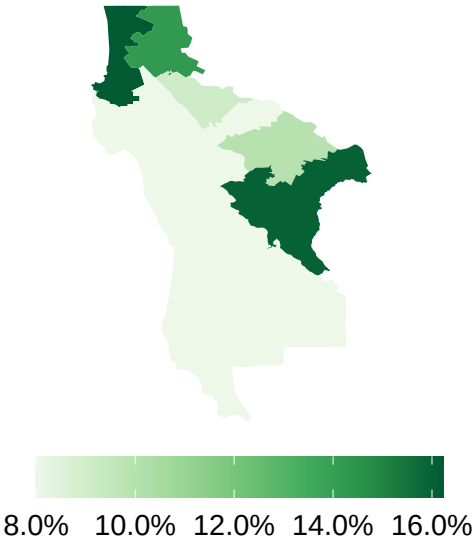
Case Rate
Per 100,000 Population



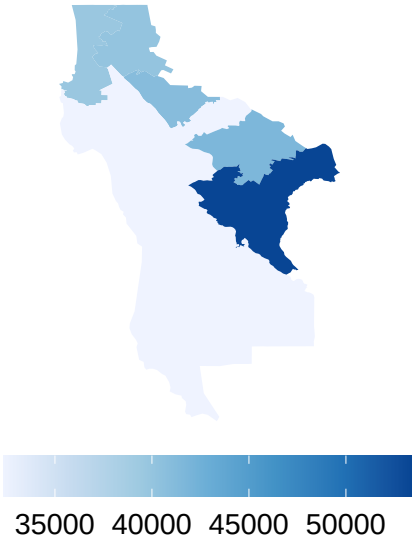
San Mateo County

Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)

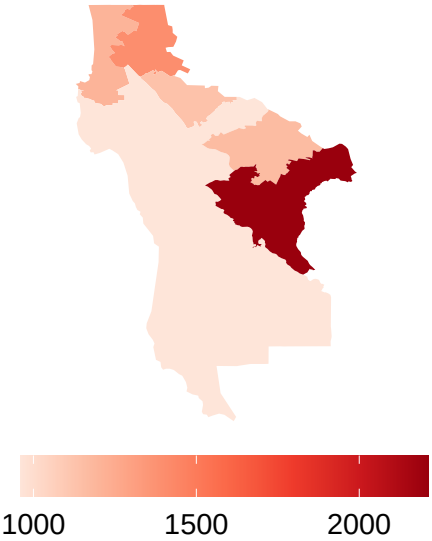
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

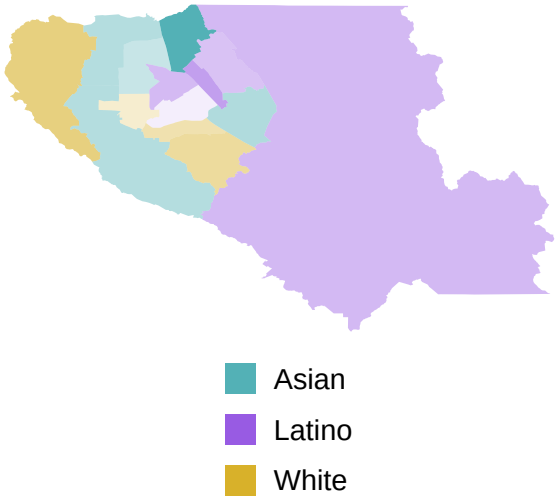


Case Rate
Per 100,000 Population

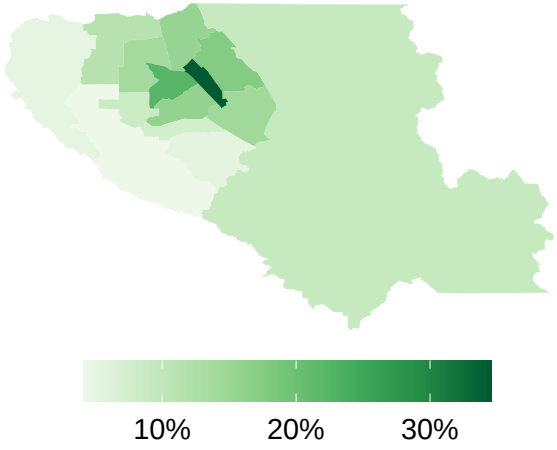


Santa Clara County

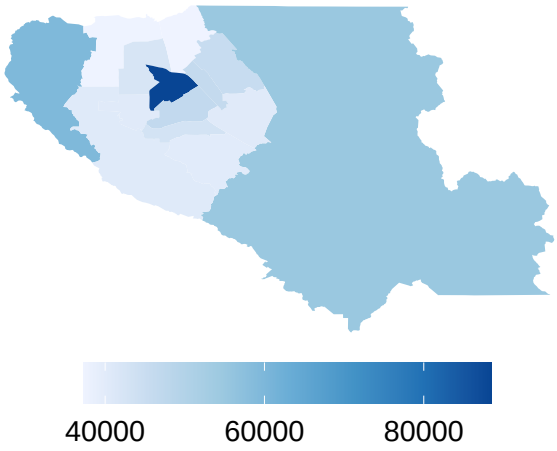
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



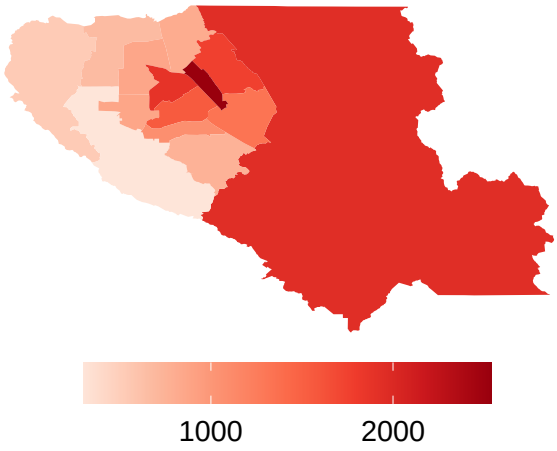
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

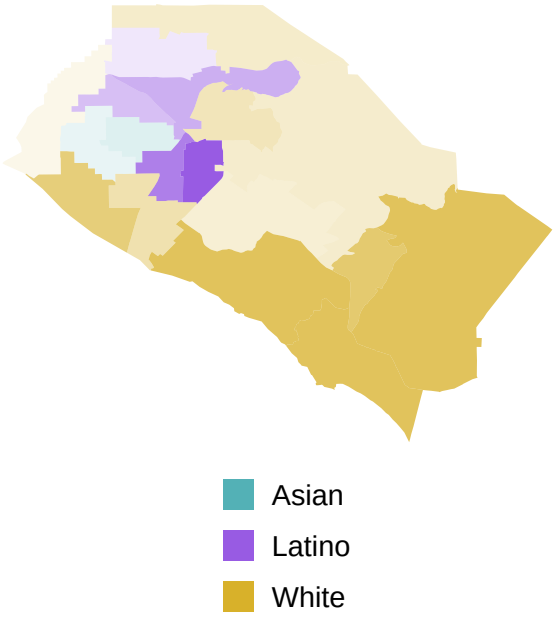


Case Rate
Per 100,000 Population

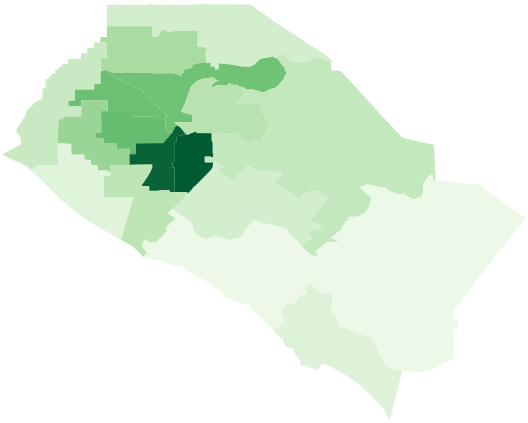


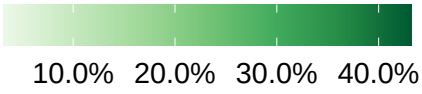
Orange County

Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)

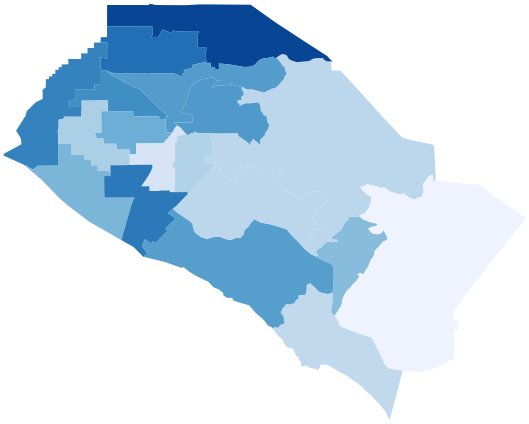


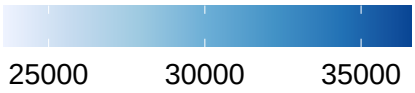
% living in crowded households
with an essential worker



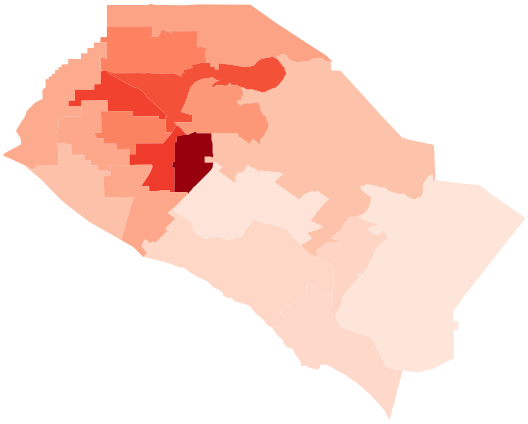


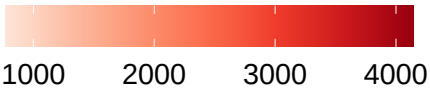
Test Rate
Per 100,000 Population





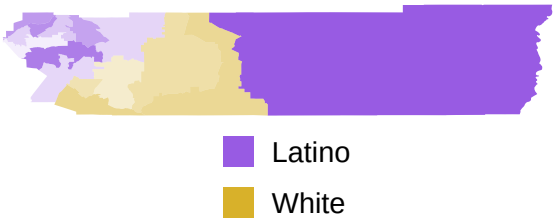
Case Rate
Per 100,000 Population





Riverside County

Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



Latino

White

% living in crowded households
with an essential worker

5.0% 10.0% 15.0%

Test Rate
Per 100,000 Population

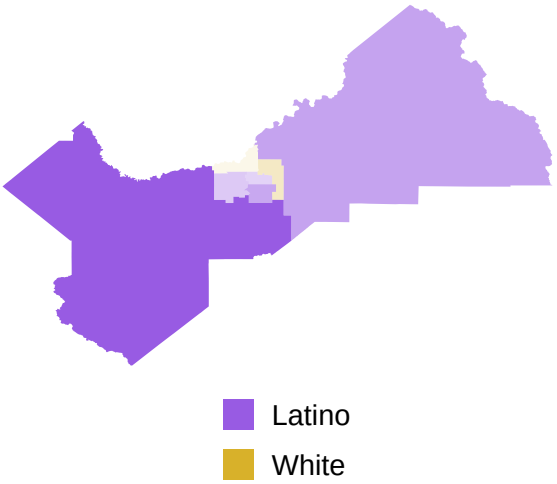
30000 35000

Case Rate
Per 100,000 Population

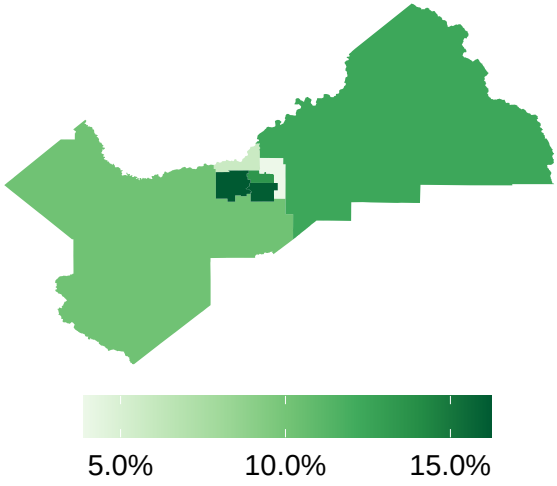
2000 3000 4000

Fresno County

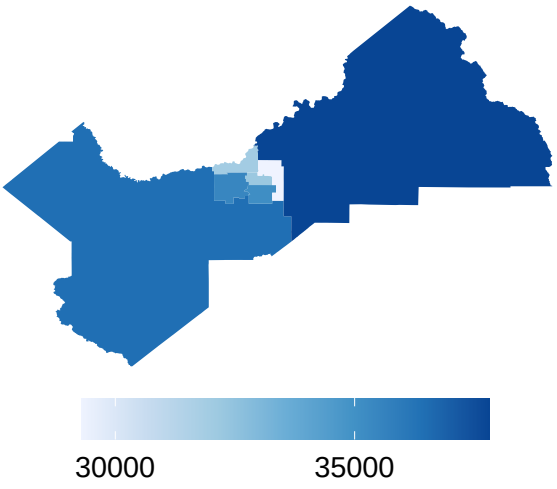
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



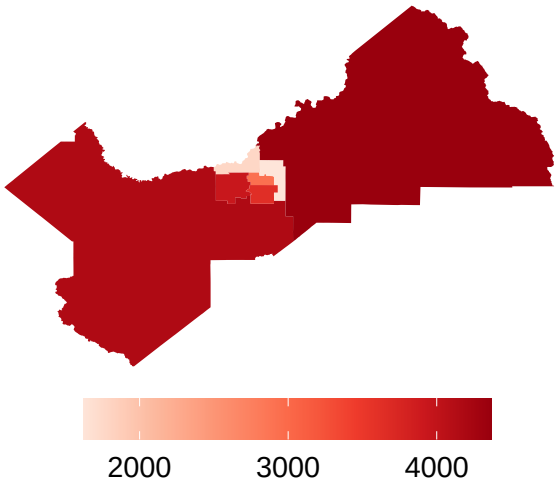
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

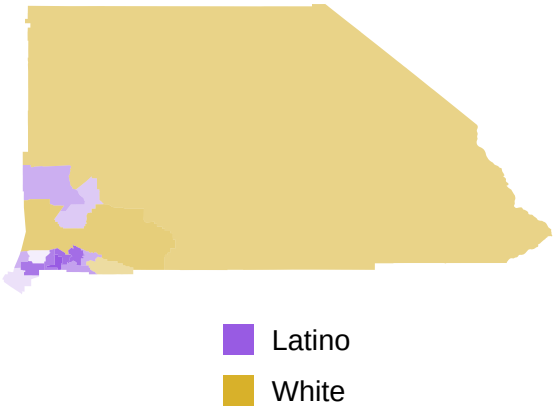


Case Rate
Per 100,000 Population

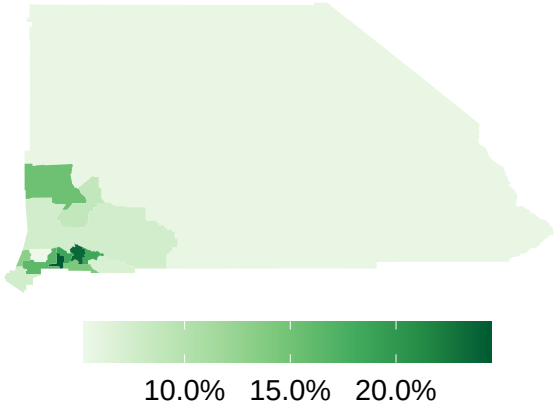


San Bernardino County

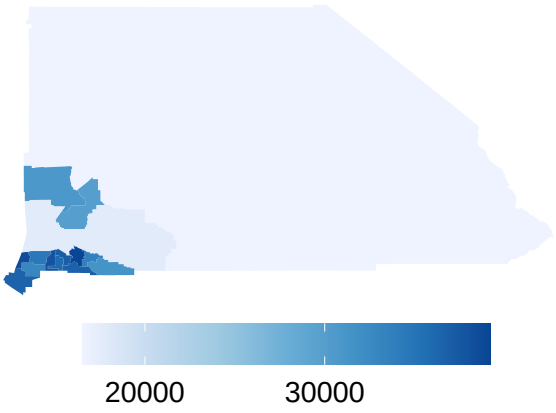
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



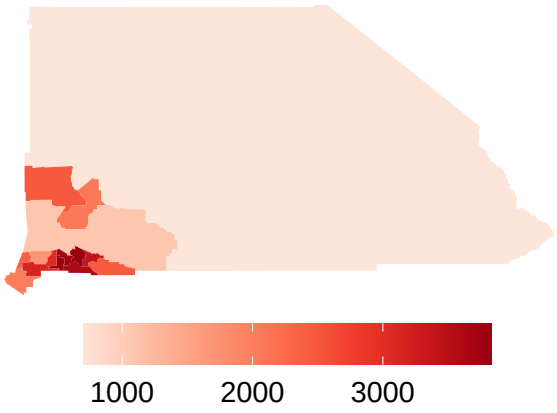
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

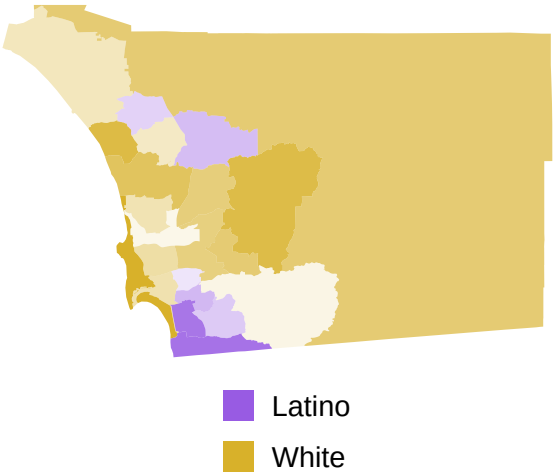


Case Rate
Per 100,000 Population

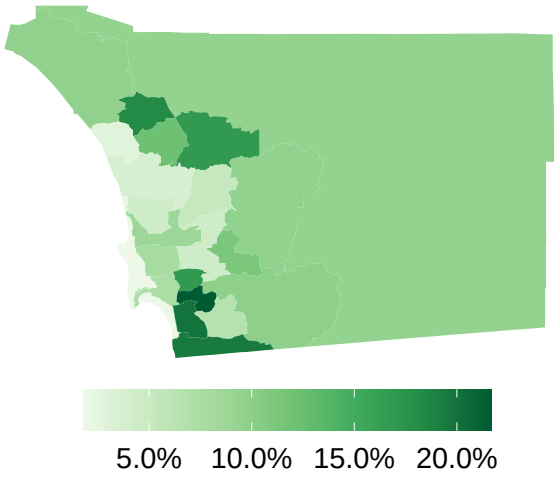


San Diego County

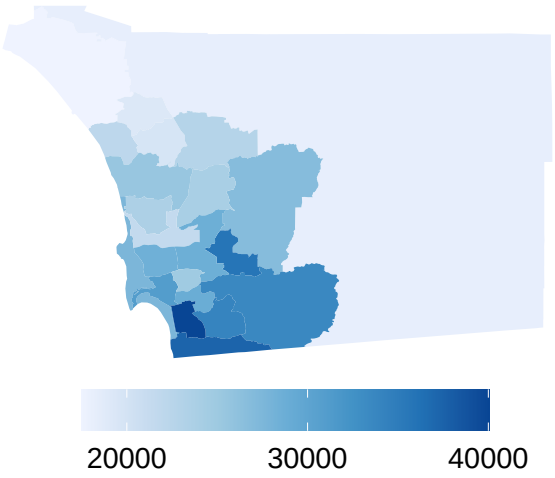
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



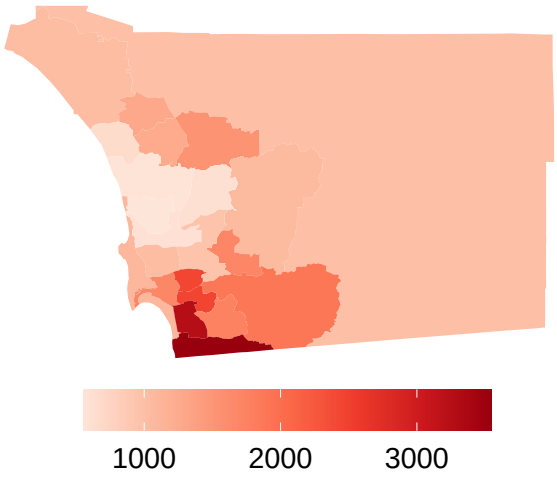
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

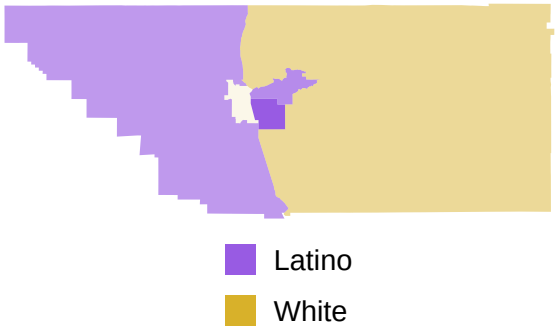


Case Rate
Per 100,000 Population



Kern County

Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)

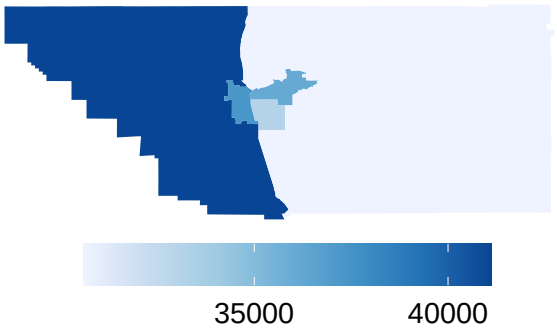


Latino

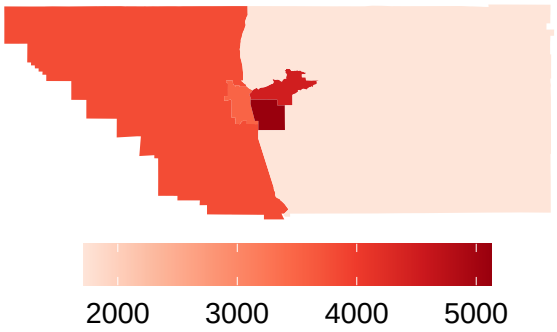
White

% living in crowded households
with an essential worker

Test Rate
Per 100,000 Population

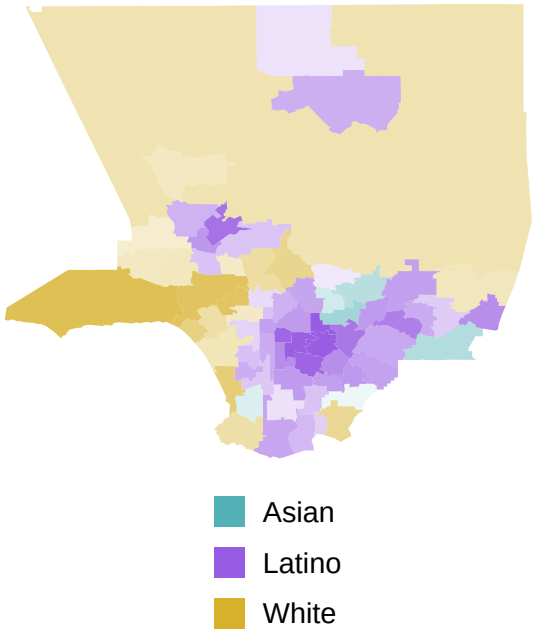


Case Rate
Per 100,000 Population

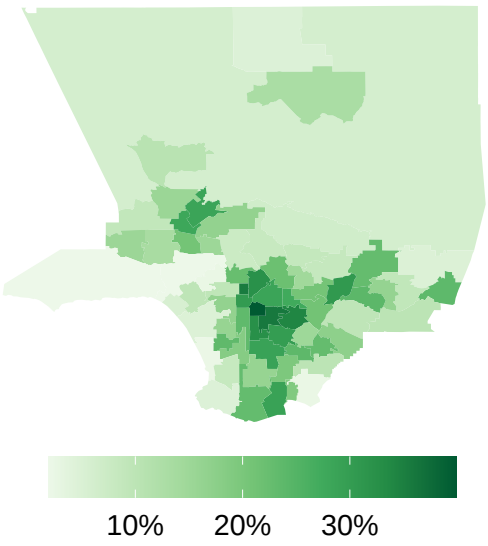


Los Angeles County

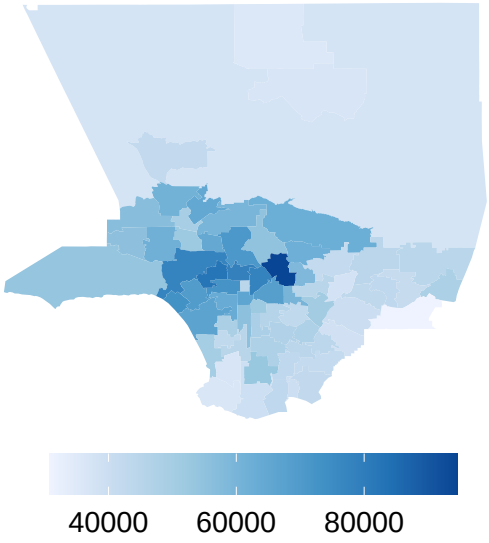
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



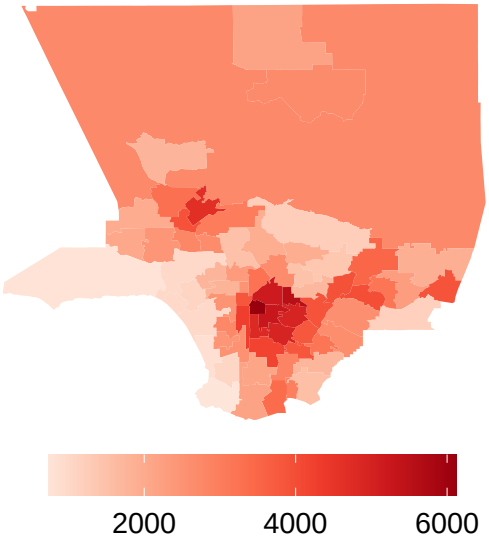
% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population

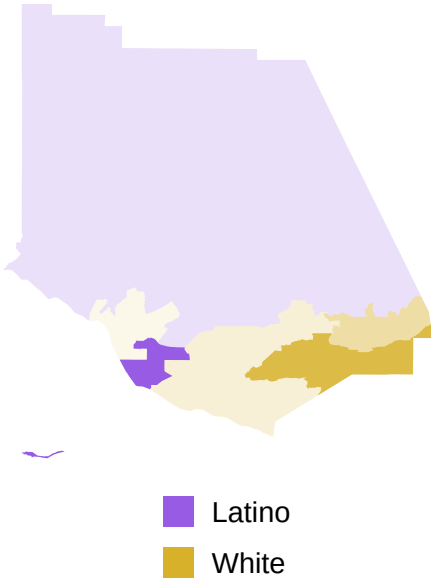


Case Rate
Per 100,000 Population

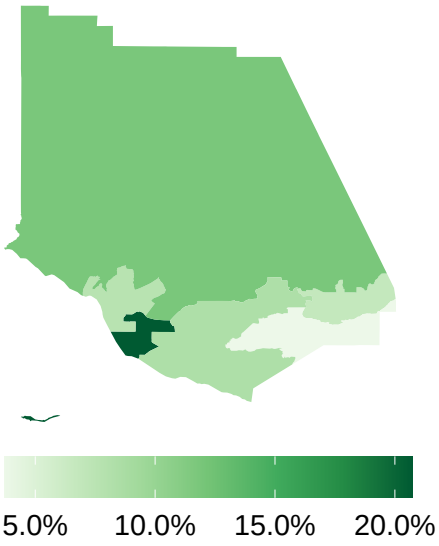


Ventura County

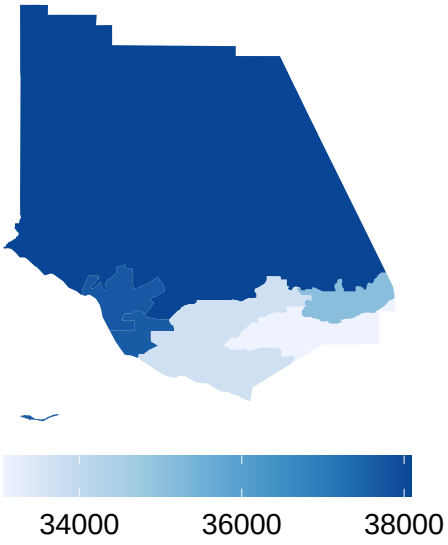
Plurality Race/Ethnicity
(Darker Color Denotes
Higher Percentage)



% living in crowded households
with an essential worker



Test Rate
Per 100,000 Population



Case Rate
Per 100,000 Population