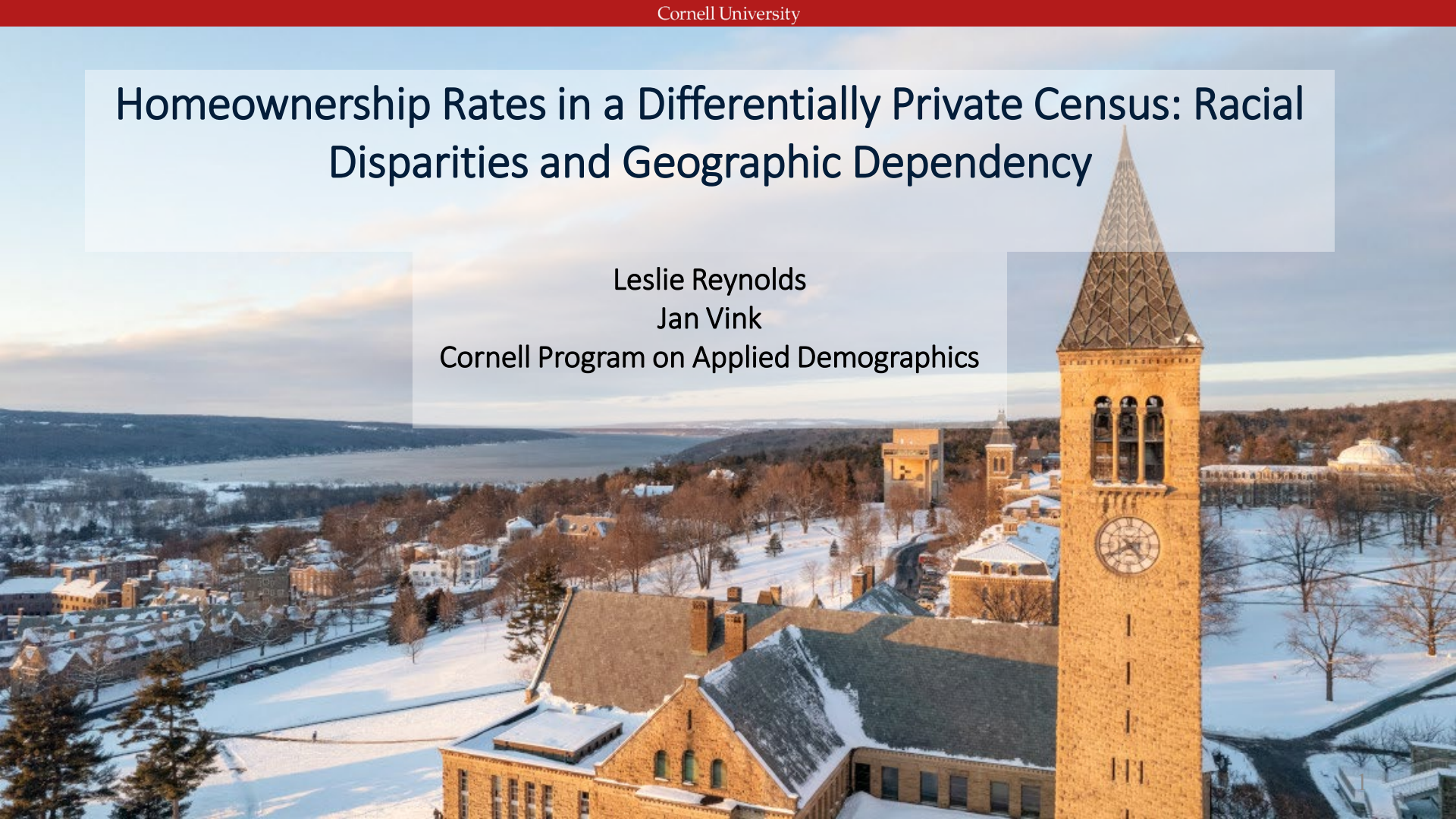# Homeownership Rates in a Differentially Private Census: Racial Disparities and Geographic Dependency
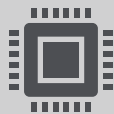
Leslie Reynolds
Jan Vink
Cornell Program on Applied Demographics

The 2020 Census Disclosure Avoidance System (DAS)

The 2020 DAS applied a variation of Differential Privacy, which injected noise according to an allocated privacy loss budget

Post-processed with TopDown algorithm so files are consistent (e.g., state population totals add to the nation's total)

Census produced a series of demonstration files (starting in 2019) that documented the development of the 2020 DAS by applying it to the 2010 Census

Allowed data users to compare original 2010 Summary File 1 with each version of the 2010 DHC

Findings from feedback on the demonstration files indicated data on small groups and geographies were more vulnerable to noise

2

Previous Analysis of the March 2022 Demo Data

People renting a home tend to have different characteristics from those who own their home (e.g., age, income level, family size)

Institutional barriers such as structural racism and wealth inequity are significant obstacles to homeownership

- Underserved and minority populations are more likely to rent than own a home (Choi et al., 2019)

Clusters of housing units by tenure tend to be scattered geographically and do not necessarily fit within the TDA of post processing

- Tenure and population totals in these areas may not be consistent with the original 2010 SF1

3

Previous Analysis of the March 2022 Demo Data (cont.)

Examined the impact of the proposed (March) DAS on household data by tenure majority area for Census tracts aggregated to different geographic levels

We Found:

▪ Broad measures of similarity between SF1 and the Demonstration data can be misleading

▪ Tracts aggregated to sub-state levels generally produced the largest errors

▪ Data for households in rental majority areas were least accurate to original 2010 SF1, owner majority areas most accurate

▪ Data on Households with children and large households in rental majority areas had largest mean "errors" and largest share of "Big errors" (MAE and MAPE >=10)

Replicated this analysis with August 2022 Demonstration Data

some improvements, but disparities between tenure areas remained

Current Analysis

- Differences between the 2010 DHC and SF1 by tenure area indicate potential changes in calculated rates of homeownership due to the DAS

- Homeownership rates are an important metric of social problems such as wealth inequity and disparate impacts of the COVID-19 Pandemic

- The Decennial Census is the primary source for producing homeownership rates (HOR) by race and ethnicity at lower levels of geography (Choi & Lee, 2021)
    - Important to understand if and how the DAS impacts homeownership rates differently across race and ethnicity

  HOR= # of households in group A that own their homes/ total occupied houses for group A

- Now, we want to investigate whether comparisons of HOR by race/ethnicity are impacted by the August 2022 version of the DAS

Current Research Questions

1. Are homeownership rates by race/ethnicity produced using the demonstration data significantly different from those produced with the original SF1?

2. Do reported disparities in homeownership rates change between the 2010 SF1 and August DHC?

3. Do differences in homeownership rates by race/ethnicity between the SF1 and DHC depend on the relative size of the racial/ethnic group?

Data and Methods

- 2010 Summary File 1 and 2010 Differential Privacy Demonstration data file (released August 2022), retrieved from IPUMS (Van Riper et al., 2022)

- Census tract-level Housing Unit and Person files, merged by geocode and aggregated to counties
    - Excluded Puerto Rico
    - 73,057 tracts to 3,143 counties

- Important to keep in mind: SF1 was impacted by swapping therefore we do not know the "ground truth", whether it is the SF1 or DHC or somewhere in between
    - Therefore, differences between the files are not inherently good or bad
    - Understanding these differences is still important
    - Aim is to provide guidance for when the 2020 DHC is released, especially for comparing the 2020 DHC with previous Censuses

7

RQ 1: Differences in Homeownership Rates (HOR) Between the DHC and SF1

Table 1: T-tests for significant percentage-point differences between file Homeownership Rates, for tracts aggregated to counties

	Mean HOR (DP, SF)	Mean PP Difference	SE	t	Direction
Total (N=3,143)	72.20%, 72.21%	-0.01	0.020	-0.32	-
White (N=3,143)	74.54%, 74.16%	**0.38*****	0.029	13.01	DP > SF
Black (N=3,053)	53.00%, 51.10%	**1.90*****	0.312	6.11	DP > SF
Asian (N=3,052)	62.98%, 62.97%	0.01	0.389	0.025	-
AIAN (N=3,118)	57.95%, 59.80%	**-1.85*****	0.306	-6.03	DP < SF
Native Hawaiian/ Pacific Islander (N=2,270)	46.38%, 52.16%	**-5.78*****	0.815	-7.09	DP < SF
Other race (N=3,103)	50.19%, 48.34%	**1.86*****	0.312	5.94	DP > SF
All Other Races[+] (N=3,140) (Other+AIAN+Asian+NHPI)	56.84%, 56.59%	0.244	0.169	1.45	-
Two or More Races (N=3,135)	57.81%, 60.37%	**-2.56*****	0.209	-12.27	DP < SF
Non-Hispanic/Latinx (N=3,143)	72.88%, 72.86%	0.03	0.020	1.35	-
Hispanic/Latinx (N=3,137)	53.87%, 54.41%	**-0.54*****	0.149	-3.60	DP < SF

p<0.001***; p<0.01**; P<0.05*
Null hypothesis of no Percentage Point difference between file homeownership rates ($HR_{DHC} - HR_{SF} = 0$)
[+]This is a test grouping to see if larger differences are attenuated by combining smaller groups

8

RQ1 (cont.) Interpreting the Differences in HOR between Files

- In both files, the same groups had the three highest homeownership rates
  - Native Hawaiian/Pacific Islander homeownership rates were lowest in the DHC, while other race HOR were lowest in the SF1

- Rank order of homeownership rates by race/ethnicity differed between files even when aggregated to the national level

Table 2: Rank Order of National Average HOR by Race/Ethnicity, 2010 DHC and SF1

Ranking	Race/Ethnicity (DHC)	Average HOR	Race/Ethnicity (SF1)	Average HOR
1	White	74.5%	White	74.2%
2	Non-Hispanic/Latinx	72.9%	Non-Hispanic/Latinx	72.9%
3	Asian	63.0%	Asian	63.0%
4	**AIAN**	58.0%	Two or More Races	60.4%
5	**Two or More Races**	57.8%	AIAN	59.8%
6	Hispanic/Latinx	53.9%	Hispanic/Latinx	54.4%
7	**Black**	53.0%	NHPI	52.2%
8	**Other Race**	50.2%	Black	51.1%
9	**NHPI**	46.4%	Other Race	48.3%

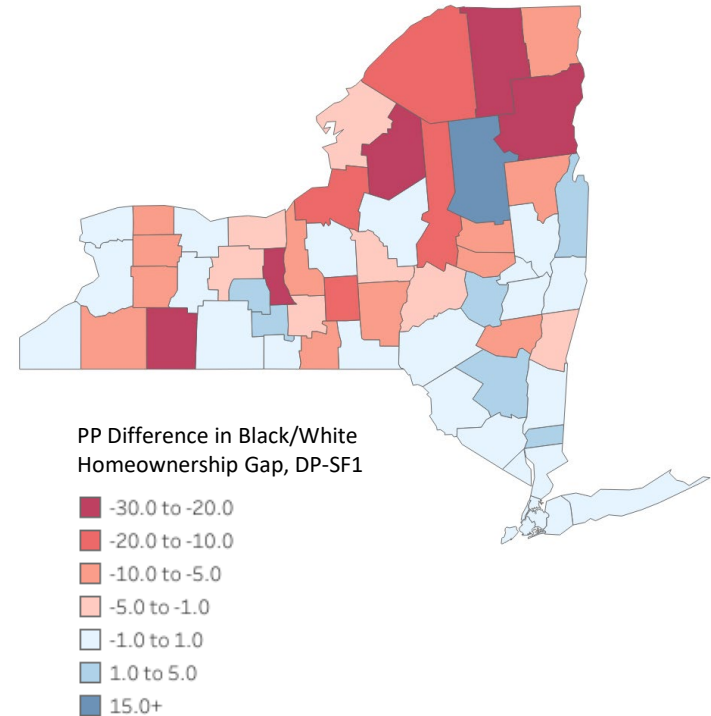Note: Groups are in bold if their HOR in the DHC differs from the SF1
AIAN= American Indian/ Alaskan Native
NHPI= Native Hawaiian/Pacific Islander

9

RQ2: File Variation in Homeownership Rate Disparities

- Black/White and Hispanic/Non-Hispanic homeownership gap sizes varied only slightly between files when aggregated to states
  - Less than half of states changed ranked order of HOR gap size between files

- More variation in homeownership gaps at the county level

- 10 NYS counties had Black-White Homeownership gaps that were over 10 percentage points smaller in the DHC than in the SF1
  - Only one county had a Black-White homeownership gap that was over 5 percentage points larger in the DHC than SF1

Figure 1: File Differences in the Black-White gap in Homeownership, New York State

PP Difference in Black/White Homeownership Gap, DP-SF1

- -30.0 to -20.0
- -20.0 to -10.0
- -10.0 to -5.0
- -5.0 to -1.0
- -1.0 to 1.0
- 1.0 to 5.0
- 15.0+

10

RQ2: File Variation in Homeownership Rate Disparities (cont.)

- County rankings of Black-White homeownership gap size could lead to different conclusions depending on the file used

  - Only 2 of 62 county Black/White homeownership gaps were ranked the same in the DHC as the SF1

- The range of gap sizes was smaller in the DHC than SF1

  - Largest 5 gaps decreased, smallest 5 gaps increased

Table 3: Top and Bottom 5 Largest County Black-White Gaps in Homeownership Rates (Percentage-Point Difference)		
Rank	SF1	DHC
1	Herkimer (46.4)	Broome (45.9)
2	Allegany (46.2)	Rensselaer (45.2)
3	Broome(45.9)	Oneida (43.2)
4	Rensselaer (45.8)	Schenectady (42.2)
5	Jefferson (45.4)	Jefferson (42.0)
58	Putnam (9.1)	Kings (9.3)
59	Bronx (7.9)	Schuyler(9.3)
60	Schuyler (7.5)	Franklin (8.8)
61	Hamilton (6.2)	Bronx (8.3)
62	Queens (0.7)	Queens (0.9)

11

RQ2: Variation in Homeownership Rate Disparities (cont.)

- File differences in the Hispanic/Non-Hispanic homeownership gap were similar to those found for Black/White homeownership

  - Smaller range of values (-14.2 to 4.3), fewer extreme values

  - Hamilton county had a Hispanic/NH gap in homeownership 14 percentage points smaller in the DHC than it was in the SF1

- County rankings of Hispanic/Non-Hispanic HOR gaps changed, but less drastically than Black/White HOR gaps

  - 15 counties maintained their ranking between the SF1 and DHC

Table 3: Top and Bottom 5 Largest County Hispanic/Non-Hispanic Gaps in Homeownership Rates (Percentage-Point Difference)

Rank	SF1	DHC
1	Genesee (44.4)	Wyoming (45.7)
2	Wyoming (43.7)	Genesee (43.9)
3	Montgomery (41.4)	Montgomery (41.0)
4	Ontario (40.7)	Chautauqua (40.1)
5	Chautauqua (39.8)	Ontario (40.0)
58	Tioga (15.7)	Franklin (11.7)
59	Franklin (12.5)	Essex (11.6)
60	Schuyler (12.0)	Schoharie (9.7)
61	Essex (10.2)	Hamilton (9.4)
62	Schoharie (8.1)	Schuyler (5.9)

# RQ3: County Variation in HOR File Differences by Group Size

Table 5: Group Size and Percentage Point Differences in County Homeownership Rates, Black Population

County Rank: Relative Size of the Black Population (Largest to smallest)	Minimum (% of population Black)	Maximum (% of population Black)	Mean Percentage-Point Difference in Black Homeownership Rates, DHC-SF1
< 1100	5.19%	85.68%	0.29
1100-2100	0.67%	5.19%	2.71
2100+	0.01%	0.67%	9.21

- When the share of the population that is Black exceeded 5.2%, the average PP difference in Black homeownership rates between the DHC and SF1 was only 0.3

- In counties whose population is less than 0.7% Black, the average PP difference in Black homeownership rates between files rose to 9.2
  - For county populations between 0.7% to 5.2% Black, the average gap was 2.7 PP

- Estimates of Black homeownership rates in counties where over 5% of the population was Black were more reliable across files

13

Conclusions

- Mean homeownership rates by race and ethnicity varied significantly across files for most groups (exceptions: Asian, Non-Hispanic population)

- Black/White and Hispanic/Non-Hispanic gaps in homeownership changed between files at the county aggregate level
    - Black/White HOR gaps were more vulnerable to changes in the DAS between 2010 and 2020

- Homeownership rate estimates for racial/ethnic groups vary in accuracy to the original SF1 depending on the relative size of the group
    - Black homeownership rates calculated for counties where over 5.2% of the population was Black had the smallest percentage point differences between files

- Before comparing rates of homeownership across Census years by race/ethnicity it will be important to evaluate potential thresholds for relative group size to create robust and comparable estimates
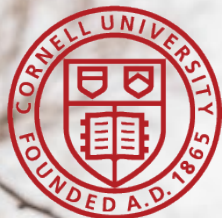
Thank you!

My information:

lr523@cornell.edu
padinfo@cornell.edu

References

- Choi, J. H., McCargo, A., Neal, M., Goodman, L., & Young, C. (2019). Explaining the Black-White homeownership gap. Washington, DC: Urban Institute. Retrieved March, 25, 2021.

- Choi, J.H., Hyojung Lee. (2021) Racial Homeownership Rates Vary across the Most Commonly Cited Datasets. When and Why Should You Use Different Ones? Washington, DC: Urban Institute. Retrieved February 3, 2023.

- Van Riper, David, Tracy Kugler, and Jonathan Schroeder. IPUMS NHGIS Privacy-Protected 2010 Census Demonstration Data, version 20220310. August 2022 Privacy Protected Data. Minneapolis, MN: IPUMS. 2020.

- 2010 Demonstration Data Product – Demographic and Housing Characteristics File (DHC) (v. 2022-08-25), Prepared by the U.S. Census Bureau, 2022