

RESEARCH

Supplementary Material

Offline Biases in Online Platforms: a Study of Diversity and Homophily in Airbnb

Victoria Koh¹, Weihua Li^{1,2}, Giacomo Livan^{1,2*} and Licia Capra¹

*Correspondence:

g.livan@ucl.ac.uk

¹Department of Computer Science, University College London, 66-72 Gower Street, WC1E 6EA London, UK
Full list of author information is available at the end of the article

1 Robustness with respect to annotation accuracy

The results presented in the main paper have been obtained by using two facial annotation softwares (Sightcorp and Betaface) in parallel. The user pictures retained in our dataset were only those for which both products provided the same annotations with a confidence higher than $0.3 \in [0, 1]$. As a robustness test of our results, we repeated our analyses on a restricted dataset limited to images for which both softwares provided the same annotations with a confidence higher than 0.5.

The summary statistics about the number of hosts, guests and pairings that we collected and annotated with confidence ≥ 0.5 for each city are reported in Table 1. The number of annotated users decreases slightly with the higher confidence threshold, but the demographic features of this subset of data remain largely unchanged with respect to those reported in the main paper (see Tables 2 and 3).

In Tables 4 and 5 we report the results obtained from the rewiring analysis on this restricted dataset on gender and race, respectively. As it can be seen, the over- and under-expression patterns we find in the gender-related pairings are exactly the same as those obtained with a lower accuracy threshold reported in the main paper. The same applies to the race-related pairings, which we find to be consistent with those obtained with a lower threshold in all but a few cases (see, e.g., Black-White pairings in Nashville’s full property rentals).

Table 1 Number of hosts, guests, and host-guest pairs annotated for each city analysed when setting the annotation confidence threshold to 0.5

City	# Hosts	# Guests	# Host-Guest Pairs
Amsterdam	2,349	68,978	70,812
Chicago	1,685	20,719	22,077
Dublin	1,033	2,590	2,757
Hong Kong	1,222	11,915	13,077
Nashville	626	1,702	2,000

As an additional robustness check of our results, we repeated our analysis on race-related pairings after an artificial manipulation of the data. Namely, we manually altered the race annotation of a randomly selected sample made of 5% of all White users in each city, changing their annotation to Black or Asian with probability $1/2$. The results obtained from the rewiring analysis on this manipulated dataset are shown in Table 6. As one might expect, this leads to a few changes with respect to the results presented in the main paper. For example, the over-expression of White-White interactions is eliminated in the case of Amsterdam. However, most

Table 2 Airbnb host and guest population by gender (F=female) in the dataset restricted to images annotated with confidence higher than 0.5

City	Full property rental		Shared property rental	
	F Host	F Guest	F Host	F Guest
Amsterdam	59%	55%	58%	58%
Chicago	62%	58%	57%	56%
Dublin	60%	59%	59%	54%
Hong Kong	55%	59%	60%	59%
Nashville	63%	66%	74%	59%

Table 3 Airbnb host population by race (W=White, A=Asian, B=Black) in the dataset restricted to images annotated with confidence higher than 0.5

City	Full property rental			Shared property rental		
	W	A	B	W	A	B
Amsterdam	94%	4%	2%	90%	9%	1%
Chicago	91%	7%	2%	89%	9%	2%
Dublin	96%	3%	1%	93%	5%	2%
Hong Kong	65%	34%	1%	59%	40%	1%
Nashville	93%	3%	4%	92%	5%	3%

Table 4 Pairings between guest-host genders (F=female; M=male) in the dataset restricted to images annotated with confidence higher than 0.5. Values in brackets represent 95% confidence level intervals obtained from the rewiring analysis, while the values to their right denote the corresponding empirically observed frequencies. Upward green (downward red) arrows highlight over-expressed (under-expressed) values.

City	Full property rental							
	FF	FF	FM	FM	MF	MF	MM	MM
AMS	[29.40; 29.45]%	30.23% ↑	[24.01; 24.05]%	23.23% ↓	[25.60; 25.64]%	24.81% ↓	[20.91; 20.95]%	21.73% ↑
CHI	[33.71; 33.79]%	34.90% ↑	[25.07; 25.15]%	23.96% ↓	[23.58; 23.66]%	22.47% ↓	[17.47; 17.55]%	18.67% ↑
DUB	[34.85; 35.11]%	37.28% ↑	[25.14; 25.40]%	22.97% ↓	[22.82; 23.07]%	20.64% ↓	[16.68; 16.94]%	19.11% ↑
HK	[33.18; 33.29]%	33.79% ↑	[24.40; 24.51]%	23.91% ↓	[24.32; 24.43]%	23.82% ↓	[17.88; 17.99]%	18.49% ↑
NAS	[41.46; 41.67]%	41.94% ↑	[22.58; 22.79]%	22.30% ↓	[23.06; 23.28]%	22.79% ↓	[12.48; 12.69]%	12.97% ↑
Shared property rental								
AMS	[28.52; 28.58]%	29.29% ↑	[20.57; 20.63]%	19.86% ↓	[29.47; 29.53]%	28.76% ↓	[21.32; 21.39]%	22.09% ↑
CHI	[31.36; 31.45]%	32.73% ↑	[24.64; 24.74]%	23.36% ↓	[24.53; 24.62]%	23.25% ↓	[19.28; 19.38]%	20.66% ↑
DUB	[30.98; 31.17]%	29.59% ↓	[26.66; 26.85]%	28.24% ↑	[22.45; 22.64]%	24.03% ↑	[19.53; 19.72]%	18.14% ↓
HK	[33.55; 33.68]%	34.22% ↑	[24.12; 24.25]%	23.58% ↓	[24.55; 24.67]%	24.00% ↓	[17.53; 17.66]%	18.20% ↑
NAS	[42.72; 43.00]%	42.84%	[30.52; 30.80]%	30.68%	[15.24; 15.51]%	15.39%	[10.97; 11.24]%	11.09%

of the homophily, heterophily, and avoidance patterns reported in the main paper do not change.

Table 5 Pairings between racial backgrounds of Airbnb guests and hosts (W=White, A=Asian, B=Black) in the dataset restricted to images annotated with confidence higher than 0.5. Values in brackets represent 95% confidence level intervals obtained from the rewiring analysis, while values below them denote the corresponding empirically observed frequencies. Upward green (downward red) arrows highlight over-expressed (under-expressed) values.

	Full property rental			Shared property rental		
AMS	W	A	B	W	A	B
W	[77.37; 77.39]%	[7.57; 7.59]%	[1.06; 1.07]%	[72.93; 72.96]%	[10.35; 10.37]%	[1.21; 1.21]%
	77.46% ↑	7.52% ↓	1.06%	73.01% ↑	10.25% ↓	1.23% ↑
A	[3.84; 3.85]%	[0.37; 0.38]%	[0.05; 0.06]%	[8.00; 8.02]%	[1.13; 1.16]%	[0.12; 0.13]%
	3.85%	0.40% ↑	0.05%	7.93% ↓	1.21% ↑	0.11% ↓
B	[1.87; 1.87]%	[0.18; 0.19]%	[0.02; 0.03]%	[1.47; 1.48]%	[0.20; 0.21]%	[0.02; 0.03]%
	1.85% ↓	0.19% ↑	0.02% ↓	1.45% ↓	0.24% ↑	0.03%
CHI	W	A	B	W	A	B
W	[72.14; 72.19]%	[9.34; 9.38]%	[1.98; 2.00]%	[67.50; 67.55]%	[12.97; 13.02]%	[2.06; 2.08]%
	72.54% ↑	9.23% ↓	1.93% ↓	67.90% ↑	12.70% ↓	2.00% ↓
A	[6.72; 6.76]%	[0.86; 0.89]%	[0.02; 0.02]%	[7.84; 7.88]%	[1.49; 1.54]%	[0.23; 0.25]%
	6.62% ↓	0.96% ↑	0.02%	7.46% ↓	1.89% ↑	0.27% ↑
B	[1.64; 1.66]%	[0.21; 0.22]%	[0.04; 0.05]%	[1.61; 1.63]%	[0.31; 0.32]%	[0.04; 0.05]%
	1.45% ↓	0.23% ↑	0.13% ↑	1.55% ↓	0.30% ↓	0.12% ↑
DUB	W	A	B	W	A	B
W	[83.04; 83.13]%	[6.88; 6.95]%	[1.56; 1.59]%	[80.46; 80.55]%	[6.31; 6.39]%	[0.76; 0.78]%
	82.99% ↓	6.98% ↑	1.45% ↓	80.51%	6.37%	0.70% ↓
A	[3.30; 3.38]%	[0.26; 0.32]%	[0.09; 0.11]%	[5.18; 5.24]%	[0.39; 0.44]%	[0.08; 0.10]%
	3.78% ↑	0.00% ↓	0.00% ↓	5.24%	0.27% ↓	0.16% ↑
B	[0.89; 0.92]%	[0.10; 0.12]%	[0.07; 0.08]%	[1.10; 1.13]%	[0.10; 0.13]%	[0.05; 0.07]%
	0.58% ↓	0.22% ↑	0.22% ↑	1.08% ↓	0.11	0.00% ↓
HK	W	A	B	W	A	B
W	[35.87; 35.99]%	[22.39; 22.50]%	[0.69; 0.71]%	[28.85; 28.96]%	[25.73; 25.84]%	[0.75; 0.77]%
	37.56% ↑	20.76% ↓	0.67% ↓	30.66% ↑	23.79% ↓	0.79% ↑
A	[20.27; 20.38]%	[12.62; 12.73]%	[0.39; 0.41]%	[19.96; 20.07]%	[17.61; 17.73]%	[0.49; 0.52]%
	18.54% ↓	14.52% ↑	0.37% ↓	18.21% ↓	19.70% ↑	0.46% ↓
B	[0.80; 0.82]%	[0.48; 0.51]%	[0.02; 0.03]%	[0.90; 0.94]%	[0.79; 0.82]%	[0.03; 0.03]%
	0.87% ↑	0.43% ↓	0.04% ↑	0.82% ↓	0.88% ↑	0.06% ↑
NAS	W	A	B	W	A	B
W	[79.27; 79.36]%	[6.63; 6.70]%	[2.04; 2.08]%	[77.23; 77.37]%	[6.29; 6.40]%	[2.13; 2.18]%
	79.16% ↓	6.57% ↓	2.12% ↑	77.83% ↑	6.67% ↑	1.51% ↓
A	[2.85; 2.90]%	[0.23; 0.28]%	[0.09; 0.11]%	[5.23; 5.35]%	[0.41; 0.50]%	[0.15; 0.18]%
	2.98% ↑	0.22% ↓	0.00% ↓	4.84% ↓	0.22% ↓	0.43% ↑
B	[3.28; 3.34]%	[0.25; 0.30]%	[0.10; 0.12]%	[1.92; 1.99]%	[0.17; 0.21]%	[0.11; 0.15]%
	3.36% ↑	0.32% ↑	0.11%	1.72% ↓	0.22% ↑	0.32% ↑

Table 6 Pairings between racial backgrounds of Airbnb guests and hosts (W=White, A=Asian, B=Black) in a manipulated dataset where a randomly selected sample of 5% of all White users are artificially annotated as either Black or Asian with probability 1/2. Values in brackets represent 95% confidence level intervals obtained from the rewiring analysis, while values below them denote the corresponding empirically observed frequencies. Upward green (downward red) arrows highlight over-expressed (under-expressed) values.

	Full property rental			Shared property rental		
AMS	W	A	B	W	A	B
W	[70.08; 70.11]% 70.10%	[9.12; 9.14] % 9.11 % ↓	[2.97; 2.98]% 1.03% ↓	[66.31; 66.35]% 66.27% ↓	[11.72; 11.75]% 11.72 % ↓	[2.91; 2.93]% 2.95%
A	[5.00; 5.01]% 5.04% ↑	[0.64; 0.66]% 0.66% ↑	[0.21; 0.22]% 0.20% ↓	[9.64; 9.67]% 9.70% ↑	[1.70; 1.73]% 1.73% ↑	[0.42; 0.44]% 0.36% ↓
B	[3.48; 3.49]% 3.49%	[0.45; 0.46]% 0.13% ↓	[0.14; 0.15]% 0.45% ↑	[2.08; 2.10]% 2.08%	[0.36; 0.37]% 0.38% ↑	[0.09; 0.09]% 0.10% ↑
CHI	W	A	B	W	A	B
W	[65.03; 65.09]% 65.41% ↑	[10.46; 10.51]% 10.41% ↓	[3.85; 3.89]% 3.85%	[59.30; 59.36]% 59.60% ↑	[13.48; 13.54]% 13.26% ↓	[3.53; 3.55]% 3.50% ↓
A	[7.77; 7.81]% 7.71 % ↓	[1.25; 1.28]% 1.32% ↑	[0.45; 0.48]% 0.42% ↓	[9.51; 9.56]% 9.16% ↓	[2.14; 2.18]% 2.55 % ↑	[0.56; 0.58]% 0.53% ↓
B	[3.34; 3.37]% 3.11% ↓	[0.53; 0.55]% 0.56% ↑	[0.19; 0.20]% 0.29% ↑	[3.79; 3.82]% 3.79% ↓	[0.87; 0.90]% 0.81% ↓	[0.22; 0.24]% 0.10% ↓
DUB	W	A	B	W	A	B
W	[74.05; 74.17]% 74.03 % ↓	[8.72; 8.83]% 8.82%	[2.79; 2.85]% 2.80 % ↓	[72.36; 72.48]% 72.23 % ↓	[8.34; 8.43]% 8.56% ↑	[3.14; 3.20]% 3.10 % ↓
A	[5.43; 5.53]% 5.88 % ↑	[0.61; 0.70]% 0.50% ↓	[0.20; 0.26]% 0.14% ↓	[6.81; 6.90]% 7.17% ↑	[0.76; 0.83]% 0.54% ↓	[0.30; 0.35]% 0.32%
B	[3.03; 3.10]% 2.65% ↓	[0.33; 0.39]% 0.50 % ↑	[0.13; 0.17]% 0.29 % ↑	[2.19; 2.24]% 2.14 % ↓	[0.24; 0.29]% 0.27%	[0.09; 0.11]% 0.11%
HK	W	A	B	W	A	B
W	[31.78; 31.89]% 33.22 % ↑	[22.03; 22.14]% 20.64% ↓	[1.52; 1.55]% 1.54%	[25.97; 26.09]% 27.38 % ↑	[24.82; 24.94]% 23.15 % ↓	[1.40; 1.43]% 1.49 % ↑
A	[19.59; 19.70]% 17.99 % ↓	[13.48; 13.59]% 15.17% ↑	[0.93; 0.96]% 0.94%	[20.06; 20.18]% 18.82% ↓	[19.24; 19.37]% 20.96% ↑	[1.07; 1.10]% 0.97% ↓
B	[2.35; 2.40]% 2.54 % ↑	[1.60; 1.65]% 1.48% ↓	[0.11; 0.12]% 0.12% ↑	[1.21; 1.24]% 1.05% ↓	[1.13; 1.17]% 1.32% ↑	[0.07; 0.08]% 0.12% ↑
NAS	W	A	B	W	A	B
W	[68.30; 68.42]% 68.09 % ↓	[8.29; 8.39]% 8.46% ↑	[5.02; 5.10]% 4.91% ↓	[68.77; 68.95]% 68.40% ↓	[7.57; 7.71]% 8.40% ↑	[3.65; 3.75]% 3.30 % ↓
A	[4.62; 4.72]% 4.91 % ↑	[0.59; 0.52]% 0.49% ↓	[0.32; 0.37]% 0.27% ↓	[6.77; 6.91]% 7.13% ↑	[0.70; 0.82]% 0.21% ↓	[0.33; 0.39]% 0.43% ↑
B	[6.20; 6.29]% 6.31 % ↑	[0.70; 0.77]% 0.65% ↓	[0.43; 0.49]% 0.65% ↑	[3.96; 4.07]% 4.04 %	[0.42; 0.50]% 0.43%	[0.23; 0.28]% 0.53% ↑

2 Robustness with respect to economic factors

As mentioned in the main paper, some of the over-/under-expressions we observe in the race-related pairings might be attributed to potential confounding factors such as users' wealth or income. Indeed, global economic inequalities often correlate with race. In particular, a White racial background typically correlates with better economic conditions. This, in turn, might partially explain some of the homophily and avoidance patterns we report in the main paper. For example, the over-expression of interactions between White hosts and guests we measured in most of the cities we analyzed could be partially (or completely) explained simply in terms of White users being on average wealthier (i.e., White hosts owning more expensive properties and White guests being able to afford more expensive stays).

We first sought to disentangle wealth and homophily via matched pair analysis. In each city we looked for pairs of White and non-White hosts with similar profiles in terms of two proxies for wealth and/or income, i.e., the number of Airbnb properties owned, and the price charged for a week-long stay at such properties. After forming such pairs (using the algorithm provided by Ref. [33] in the main paper), we measure the rate of interaction with White guests across the two groups, and run a *t*-test on the two rates against a null hypothesis of equal rates of interaction.

The results of this analysis are reported in Table 7. As it can be seen, we found statistically significant differences in the rates of interaction with White guests across the two groups only in the case of Hong Kong (regardless of the type of property) and Chicago's full property rentals. This is due to the imbalance in the data, since in all cities we analyze White hosts are the vast majority, which makes it rather hard to find large enough numbers of pairs with non-White hosts. As a matter of fact, the most significant results were found in Hong Kong, which is by far the most diverse city among the ones we analyzed.

Table 7 Results of the matched pair analysis

City	Property	White/White	non-White/White	Pairs	Stays	<i>p</i> -value
Amsterdam	Shared	82.31%	83.29%	56	2301	0.287
Amsterdam	Full	84.88%	85.02%	187	4464	0.833
Chicago	Shared	74.42%	75.64%	141	2905	0.303
Chicago	Full	81.92%	79.45%	148	1792	0.0451*
Dublin	Shared	89.08%	90.87%	69	403	0.475
Dublin	Full	91.89%	88.80%	33	111	0.422
Hong Kong	Shared	56.21%	47.67%	217	3535	< 0.001***
Hong Kong	Full	60.31%	53.00%	260	3051	< 0.001***
Nashville	Shared	84.03%	82.08%	26	119	0.698
Nashville	Full	85.28%	88.64%	51	598	0.233

In order to overcome this issue, we sought to control for the hosts' wealth / income by removing from each network the hosts belonging to the top and bottom third of the distribution of prices charged for a week-long stay. This left us with bipartite sub-networks made exclusively of owners of middle-range properties and their guests. We report the results of our network rewiring analysis obtained on such sub-networks in Table 8. Once again, we find the results to be very much in line with those reported in the main paper, with a few exceptions (most notably, interactions involving White guests in Dublin).

Table 8 Pairings between racial backgrounds of Airbnb guests and hosts (W=White, A=Asian, B=Black) in the sub-networks obtained by removing all hosts belonging to the top and bottom thirds of the distribution of prices charged for a week-long stay. Values in brackets represent 95% confidence level intervals obtained from the rewiring analysis, while values below them denote the corresponding empirically observed frequencies. Upward green (downward red) arrows highlight over-expressed (under-expressed) values.

	Full property rental			Shared property rental		
AMS	W	A	B	W	A	B
W	[80.30; 80.33] % 80.33% ↑	[8.14; 8.16] % 8.09% ↓	[1.02; 1.03] % 1.03%	[68.30; 68.37] % 68.75% ↑	[7.74; 7.80] % 7.69% ↓	[1.10; 1.13] % 1.06% ↓
A	[4.37; 4.40] % 4.34% ↓	[0.43; 0.45] % 0.52% ↑	[0.05; 0.06] % 0.06%	[2.64; 2.67] % 2.49% ↓	[0.29; 0.31] % 0.42% ↑	[0.04; 0.05] % 0.03% ↓
B	[1.34; 1.35] % 1.38% ↑	[0.13; 0.14] % 0.13%	[0.02; 0.02] % 0.00% ↓	[0.85; 0.87] % 0.84% ↓	[0.10; 0.12] % 0.13% ↑	[0.02; 0.03] % 0.01% ↓
CHI	W	A	B	W	A	B
W	[74.41; 74.48] % 74.91% ↑	[9.59; 9.65] % 9.56% ↓	[2.14; 2.17] % 2.04% ↓	[69.20; 69.30] % 69.49% ↑	[11.16; 11.25] % 11.07% ↓	[1.59; 1.62] % 1.46% ↓
A	[4.57; 4.62] % 4.44% ↓	[0.56; 0.61] % 0.77% ↑	[0.01; 0.01] % 0.01%	[7.70; 7.78] % 7.32% ↓	[1.25; 1.32] % 1.60% ↑	[0.17; 0.19] % 0.31% ↑
B	[2.24; 2.28] % 1.98% ↓	[0.27; 0.30] % 0.19% ↓	[0.06; 0.07] % 0.23% ↑	[1.90; 1.93] % 1.87% ↓	[0.29; 0.32] % 0.27% ↓	[0.05; 0.07] % 0.10% ↑
DUB	W	A	B	W	A	B
W	[83.15; 83.30] % 83.29%	[5.28; 5.40] % 5.22% ↓	[1.89; 1.97] % 2.09 % ↑	[81.37; 81.52] % 81.45%	[6.44; 6.55] % 6.18% ↓	[0.84; 0.87] % 0.93% ↑
A	[3.96; 4.07] % 4.18 % ↑	[0.33; 0.41] % 0.26% ↓	[0.25; 0.29] % 0.00% ↓	[2.85; 2.93] % 3.09% ↑	[0.26; 0.33] % 0.15% ↓	[0.15; 0.15] % 0.00% ↓
B	[0.00; 0.00] % 0.00%	[0.00; 0.00] % 0.00 %	[0.00; 0.00] % 0.00 %	[0.54; 0.57] % 0.46% ↓	[0.15; 0.18] % 0.15	[0.00; 0.00] % 0.00%
HK	W	A	B	W	A	B
W	[38.27; 38.42] % 39.35% ↑	[21.64; 21.78] % 20.22% ↓	[0.85; 0.90] % 0.98% ↑	[28.40; 28.59] % 29.56% ↑	[23.20; 23.38] % 21.77% ↓	[0.74; 0.78] % 0.91% ↑
A	[19.55; 19.70] % 18.49 % ↓	[11.12; 11.25] % 12.68% ↑	[0.45; 0.49] % 0.36% ↓	[22.61; 22.81] % 21.25% ↓	[18.31; 18.50] % 20.20% ↑	[0.58; 0.63] % 0.48% ↓
B	[0.91; 0.95] % 1.04 % ↑	[0.51; 0.55] % 0.42% ↓	[0.04; 0.05] % 0.03% ↓	[0.04; 0.04] % 0.00% ↓	[0.04; 0.04] % 0.04%	[0.00; 0.00] % 0.00%
NAS	W	A	B	W	A	B
W	[79.63; 79.75] % 79.35% ↓	[8.24; 8.34] % 8.54% ↑	[1.40; 1.44] % 1.41%	[73.89; 74.29] % 74.31% ↑	[3.48; 3.72] % 3.47% ↓	[1.17; 1.28] % 0.69 % ↓
A	[2.94; 3.02] % 3.24 % ↑	[0.31; 0.38] % 0.11% ↓	[0.11; 0.16] % 0.00% ↓	[10.13; 10.46] % 9.72% ↓	[0.81; 1.03] % 0.69% ↓	[0.69; 0.69] % 0.69%
B	[2.39; 2.46] % 2.59 % ↑	[0.26; 0.32] % 0.22% ↓	[0.10; 0.12] % 0.00% ↓	[0.00; 0.00] % 0.00 %	[0.00; 0.00] % 0.00%	[0.00; 0.00] % 0.00%

Author details

¹Department of Computer Science, University College London, 66-72 Gower Street, WC1E 6EA London, UK.

²Systemic Risk Centre, London School of Economics and Political Science, Houghton Street, WC2A 2AE London, UK.

References