City Population, Majority Group Size, and Residential Segregation Drive Implicit Racial Biases in U.S. Cities

Andrew J. Stier[1*], Sina Sajjadi[2,3], Fariba Karimi[2],
Luís M.A. Bettencourt[4,5], Marc G. Berman[1,6*]

[1]Department of Psychology, University of Chicago, Chicago, IL, USA
[2]Complexity Science Hub, Vienna, Austria
[3]Central European University, Vienna, Austria
[4]Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA
[5]Mansueto Institute for Urban Innovation, University of Chicago, Chicago, IL, USA
[6]The University of Chicago Neuroscience Institute, University of Chicago, Chicago, IL, USA

[*]To whom correspondence should be addressed; E-mail:
andrewstier@uchicago.edu or bermanm@uchicago.edu.

Implicit biases, expressed as differential treatment towards out-group members, are pervasive in human societies. These biases are often racial or ethnic in nature and create disparities and inequities across many aspects of life. Recent research has revealed that implicit biases are, for the most part, driven by social contexts and local histories. However, it has remained unclear how and if the regular ways in which human societies self-organize in cities produce systematic variation in implicit bias strength. Here we leverage extensions of the mathematical models of urban scaling theory to predict and test between-city differences in implicit racial biases. Our model comprehensively links scales of organization from city-wide infrastructure to individual psychology to quanti-

tatively predict that cities that are (1) more populous, (2) more diverse, and (3) less segregated have lower levels of implicit biases. We find broad empirical support for each of these predictions in U.S. cities for data spanning a decade of racial implicit association tests from millions of individuals. We conclude that the organization of cities strongly drives the strength of implicit racial biases and provides potential systematic intervention targets for the development and planning of more equitable societies.

Introduction

Cities are organized in surprisingly regular ways (*1–3*), which drive and constrain social interactions similarly across cultures and time (*4–6*). However, there are many factors beyond the built-space geometry (*2, 3*) of cities that modulate urban social interactions. Among these, implicit biases towards out-group members are one of the most universal (*7*). Implicit biases refer to differential treatment of individuals who belong to out-groups, in ways that are automatic. These biases pose major barriers to equity and, in particular, implicit *racial* biases have been associated with disparities across essentially all aspects of life, including medical care (*8*), scholastic performance (*9*), employment (*10*), policing (*11, 12*), mental health outcomes (*13*), and physical health (*14*). If city organization and structure contribute meaningfully to these biases, there may be ways to leverage such regularities to systematically intervene and design for less biased urban areas. Despite the universality of implicit racial and ethnic biases in human societies and their well-documented detrimental effects, there have been no investigations of how the organization of people in cities may systematically influence them.

Early investigations of the origins of implicit racial biases revealed that they develop early in life (*15, 16*), are stable into adulthood, and are less prevalent in schools with more diverse populations (*16*). Neurobiological evidence complemented these findings and showed that individuals with lower levels of bias process out-group stimuli more automatically. In particular, lower levels of implicit biases are associated with more automatic processing and less activation of a network of brain areas related to social context (*17–20*). These observations suggested that early childhood exposure to diverse individuals is critical for building out-group expertise and locking-in low levels of implicit biases (*21–23*).

More recent work has demonstrated, however, that interventions with older children and adults that increase exposure to out-group individuals also reduce implicit biases, although these effects wear off if the intervention is not continued (*24–26*). This suggests that individuals' biases likely reflect ongoing predictions about their social environment (*27,28*), and consequently, that consistent population averages of implicit biases (*29*) are the result of consistent social contexts. Thus, earlier findings of stable implicit biases throughout adulthood likely reflect, in fact, not stable individual cognitive biases but instead the stability of social environments (*27–30*).

For example, the effects of slavery and associated racial segregation in the United States (U.S.) on social context and network structure have been enduring. Areas in the U.S. with larger slave populations in 1860 have higher current levels of implicit racial biases today (*30*). This example demonstrates one way in which longstanding structural influences on social contexts (e.g., racism) may drive implicit biases and perpetuate them across generations. Given the strong influence of city organization on urban social interactions and contexts (*1, 2, 31*), it is natural to ask if there are general ways in which urban environments influence and shape implicit biases.

Results

Inter-Group Interactions in Cities

We start our analysis of urban composition from the point of view of urban scaling theory (*1,2*). Its mathematical models describe cities as social networks enabled and structured by self-consistent hierarchical infrastructure networks. In this type of model, cities arise as the result of balancing the spatial costs of housing and the transportation of goods and people with the benefits of facilitating social interactions over cities' infrastructure networks (*1, 2*). These models derive average properties of cities as a function of their population size, N, as scale-invariant scaling laws (*1, 2*). For example, in the case of average per-capita social interactions, k, the scaling law takes the form of $k \sim N^{\delta}$, where $\delta = \frac{1}{6}$.

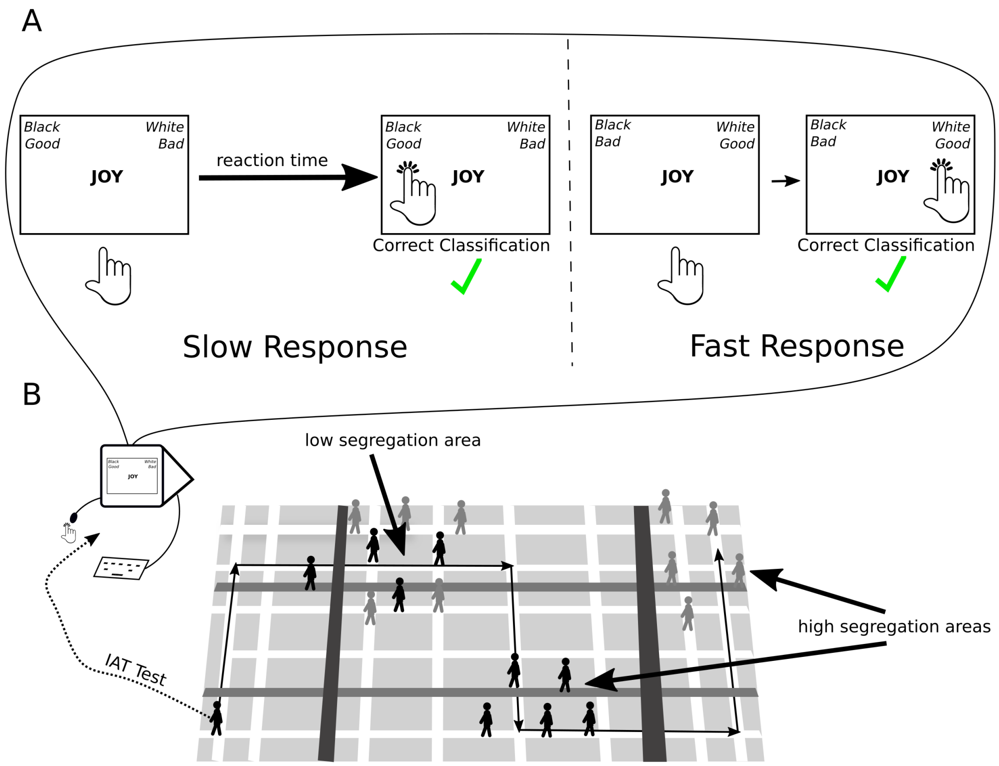


Figure 1: A) The Implicit Association Test measures implicit racial biases as a relative difference in reaction times between different pairings of word and face categories. B) We model implicit racial biases in cities as a cumulative exposure process to out-group individuals shaped by city population size, demographic diversity, and residential racial segregation.

In the simplest models of urban scaling theory, all urban inhabitants are taken to be equally likely to interact (i.e, there is homogeneous mixing) and all inhabitants are treated, in this sense, identically. In our related work, we developed modifications of these models to account for individuals belonging to distinct groups and for the fact that their connections may be biased by group identities, such that individuals may interact less often with out-group individuals and more often with their in-group (*32*). This translates into groups that may show homophilic and heterophobic interaction tendencies. In developing these models, our focus was on understanding how homophily/heterophobia and group sizes impact emergent socio-economic outputs in cities as a result of the inhibition of a number of interactions across individuals of different racial and ethnic groups. However, here, we focus more directly on what this model can reveal about systematic variations in inter-group interactions and subsequent consequences for implicit biases.

The model of heterogeneous group interaction describes the number of per-capita interactions, k_i in city i, on average, as:

$$k_i \sim N_i^{\delta} \left[\sum_{g=1}^{G} (\frac{N_{g,i}}{N_i})^2 (1 + h_{g,i}^{hom}) + \sum_{g=1}^{G} \sum_{j \neq g} \frac{N_{g,i}}{N_i} \frac{N_{j,i}}{N_i} (1 - h_{g,i}^{het}) \right] \tag{1}$$

Here, g indexes the distinct groups in cities, $h_{g,i}^{het}$ and $h_{g,i}^{hom}$ are the heterophobia and homophily values of group g in city i, and $N_{g,i}$ is the population of group g in city i. In this model, individuals from group g in city i interact with out-group individuals with a relative rate $1 - h_{g,i}^{het}$ and with in-group of $1 + h_{g,i}^{hom}$ (*32*). In addition, we have made the assumption that each group avoids all other groups similarly so that there are no unique heterophobia effects between pairs of groups (*32*).

The first term of Equation 1 is the typical scaling law (*1, 2, 4*). The second term has two components, each representing fractions of the total number of possible social interactions, N^2. The first of these captures social interactions which occur within groups, on average: $k_{\text{within},i} \sim N_i^{\delta} \cdot \sum_{g=1}^{G} (\frac{N_{g,i}}{N_i})^2 (1 + h_{g,i}^{hom})$. The second term captures social interactions which occur between groups, on average: $k_{\text{inter},i} \sim N_i^{\delta} \cdot \sum_{g=1}^{G} \sum_{j \neq g} \frac{N_{g,i}}{N_i} \frac{N_{j,i}}{N_i} (1 - h_{g,i}^{het})$. Since previous research has qualitatively indicated that inter-group interactions shape implicit racial biases (*16,* *24, 27, 28, 33–37*), we focus on this term to build our model.

In order to explicitly connect the quantity of inter-group interactions in cities to implicit bias levels, an additional step is required to translate from inter-group interactions to levels of implicit biases (*1*). Previous research has suggested that this relationship is positive – more inter-group interactions are associated with lower implicit bias levels (*16, 24, 27, 28, 33–37*). In addition, neurobiological studies provide evidence that individuals with lower levels of bias engage in more automatic processing of out-group stimuli, indicating greater expertise (*17–20*).

A common feature of such expertise-based learning is decreasing marginal returns to exposure, which is often formalized in a learning curve (*38–41*). Learning curves describe the relationship between costs and expertise across diverse individual or group tasks such as motor learning (*39*), sequence learning (*40*), solar panel construction (*41*), and cigar rolling (*38*). Typically, these learning curves are described by power-laws of the form $cost \sim n^{-\alpha}$, where n is the number of learning instances, and $1 > \alpha > 0$ determines the speed of learning (or learning rate, $\alpha = -d \ln cost / d \ln n$), with larger values of α implying faster learning. Such learning curves are a natural modeling choice to couple inter-group interactions and implicit bias levels since our measure of implicit bias, b, can be interpreted as a cognitive processing cost: b is a relative difference in reaction times when pairing photographs of white and black faces with positive and negative words, see Materials and Methods. Thus, decreasing b can be seen in this context as learning that increases social performance in a diverse population, and such learning is the implied result of greater levels of exposure (interactions) to out-group individuals.

With the additional assumption that coupling strength and direction do not vary between different pairs of groups or across interaction types (e.g., friendship, employment, acquaintance, etc) (*1, 2*), we expect measured bias levels to follow a learning curve of $b_i \sim k_{\text{inter},i}^{-\alpha}$ and therefore, we predict larger cities systematically have lower levels of bias according to:

$$b_i \sim N_i^{-\delta\alpha} \cdot \left[\sum_{g=1}^{G} \sum_{j \neq g} \frac{N_{g,i}}{N_i} \frac{N_{j,i}}{N_i} (1 - h_{g,i}^{het}) \right]^{-\alpha} \tag{2}$$

In the presence of heterophobia, it is interesting to consider the case of cities with only two distinct groups. This approximation is particularly relevant to the measure of implicit racial bias we employ here, which explicitly contrasts white and black racial groups. In this case, the scaling law for implicit racial biases simplifies to (see Supplementary Text):

$$b_i \sim N_i^{-\delta\alpha} \cdot \left[\frac{N_{1,i}}{N_i} - (\frac{N_{1,i}}{N_i})^2 \right]^{-\alpha} \cdot (2 - h_{1,i}^{het} - h_{2,i}^{het})^{-\alpha} \tag{3}$$

Equation 3 can be understood in terms of three multiplicative terms: a scaling law, a majority group size adjustment, and a heterophobia adjustment. Inter-group interactions drop dramatically as the majority group size increases and less dramatically as the heterophobia values of the groups increase (see Methods). In practice, since some cities are not very diverse ($\frac{N_1}{N} \sim 1$) and heterophobia values are small (Supplementary Figure 1), the majority group size adjustment is expected to play a much larger role than the heterophobia adjustment in determining the average number of inter-group interactions and in driving subsequent implicit biases.

In addition, Equation 3 also predicts that the logarithms of the majority group size adjustment and the heterophobia adjustment should be negatively and linearly related to the logarithm of implicit bias, b. These two adjustment terms capture deviations from the mean-field scaling law ($b \sim N^{-\delta\alpha}$) due to the specific characteristics of each given city. In summary, the model predicts that larger, more diverse, and less heterophobic cities have lower average levels of implicit biases.

Finally, the model suggests that deviations of the scaling exponent away from $\delta = \frac{1}{6}$ and the magnitude of the majority group size effect can provide empirical estimates of the learning rate, α, which characterizes the coupling between inter-group interactions and implicit racial biases. Since heterophobia values are not directly observed (see Materials and Methods), we cannot obtain a direct estimate of α from the third term of Equation 3. In addition, we note that there may be other sources of deviations from the expected scaling exponent of $\delta = \frac{1}{6}$ including top-down hierarchical constraints on inter-group interactions (*42*), growth rate fluctuations, and other higher-order effects (*43*), which may contribute to differences in independent estimates of α calculated from the first and the second terms of Equation 3.

Empirical Tests of the Urban Scaling Model of Inter-Group Bias

We next test the three predictions of our model: (1) that implicit biases systematically decrease with city size via a scaling law of $b \sim N^{-\delta\alpha}$, (2) that cities with larger majority group sizes have higher levels of implicit biases, and, (3) that less heterophobic cities have lower levels of implicit biases.

We used data from the racial Implicit Association Test (IAT) to quantify the level of implicit racial bias in U.S. cities for each year in 2010-2020 (*44*). The racial IAT measures the difference in response times when subjects pair images of white versus black faces with positive or negative words. We linked average IAT bias scores from approximately 2.7 million individuals in combined statistical areas (CBSAs) with racial demographics and population data from the U.S. Census to test our predictions. We note that CBSAs are functional definitions that capture the spatiotemporally extended social networks of cities and include, in the same unit, where people live, socialize, and work (*45*). We measured heterophobia values, h_i^{het}, as linearly dependent on residential racial segregation calculated from racial demographics in census tracts (small areas of $\sim 4,000$ inhabitants). We repeated this statistical analysis across four different measures of residential racial segregation, as in our related work (*32*). We find that across all years and measures of residential racial segregation, larger cities have lower levels of implicit racial biases, in line with Equation 3 (Figure 2A, Supplementary Table 1).

In addition, larger majority group sizes and higher levels of residential racial segregation are significantly related to scaling deviations (Supplementary Table 1) and associated with higher average IAT scores, in line with Equation 3. We note that when analyzing single years of

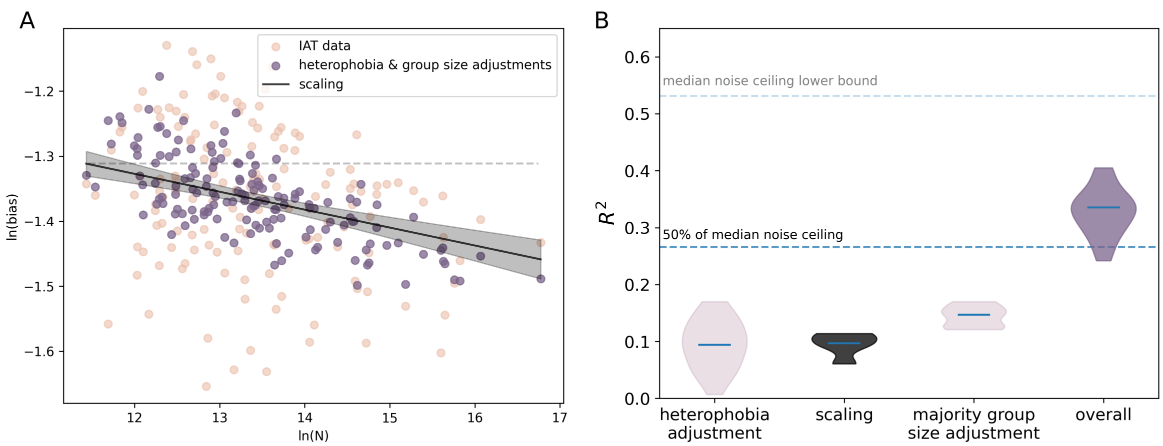


Figure 2: A) Scaling relationship, majority group size adjustment, and heterophobia adjustment for IAT data from 2020 in 149 cities with > 500 IAT responses per city. The shaded region is the 95% confidence interval for the scaling relationship. For visualization purposes, the heterophobia adjustment shown in this figure were estimated using only the mean deviation segregation measure. Results were similar with cutoffs of > 250 and > 1000 IAT responses per city and for other measures of segregation (Supplementary Tables 37-44). B) Variance explained (R^2) by the heterophobia adjustment (measured via residential racial segregation), majority group size adjustment, and scaling relationship. Data are shown for 2016-2020. Medians are shown by a horizontal line and have values of 0.094, 0.097, 0.147, and 0.346, respectively. Variance explained by the heterophobia adjustment is from all four models with different segregation measures. Noise ceiling estimates were obtained by computing correlations of bias levels between split halves of IAT participants within cities.

data before 2015, residential racial segregation is not significantly related to scaling deviations for some segregation measures. However, this is likely due to much lower sample sizes in those years resulting in fewer cities with available data and smaller fractions of city populations represented (Supplementary Table 2).

Further, the city size scaling, majority group size, and residential racial segregation effects are predictive of individual IAT responses when controlling for race, birth-sex, and educational attainment (Supplementary Tables 3-13). This suggests that these large-scale structural determinants of implicit racial biases are relevant for individuals' levels of bias. In other words, city-wide organizational and structural characteristics influence individual implicit biases despite the diversity of local social environments that any individual urban inhabitant might encounter.

Along these lines, other research has identified environmental variables related to area deprivation associated with inter-city variance in implicit racial bias (*46*). However, with our model, we find that measures of area deprivation independently explain only a small portion of the variance in inter-city differences above and beyond the three structural factors we identify here (Supplementary Tables 14-17). This suggests that the area deprivation variables identified previously actually capture a combination of city population, segregation, and majority group size and that there are other factors, for example, segregated mixing in ambient populations (*47*), that may explain the remaining inter-city variance in implicit biases.

In addition, we observe that for 2015-2020, systematic variations in city size, majority group size, and heterophobia account for a median of 33.6% (with a range of [24.2%, 40.5%]) of the variance in implicit racial bias across cities (and all four segregation measures), which is equivalent to a correlation of $r \sim 0.58$ (range of $r \sim [0.49, 0.64]$, Figure 2B, Supplementary Tables 18-24). Estimates of the noise ceiling (*48*) suggest that these three structural factors may capture a majority of the variance than can be accounted for given the reliability (*49*) of the IAT measure (Supplementary Tables 25-36). As expected, based on the fact that many U.S. cities are not diverse, majority group size accounts for more between-city variance in implicit bias than residential racial segregation (Figure 2B).

Finally, we compared estimates of the learning rate, α to previously conducted experimental interventions (*25, 26*) designed to simulate inter-group contact. The two independent estimates of α, from the scaling exponent and the majority group size adjustment (see Materials and Methods), are convergent and consistent (Figure 3). This need not have been the case and this convergence of estimates provides empirical support for a shared mechanism (namely a learning curve as a function of out-group exposure) coupling city population and majority group size to implicit bias levels. These empirical estimates of the learning rate are also consistent with experimental interventions – in which simulated inter-group contact is overwhelmingly positive and occurs immediately before bias measurements – that provide an upper bound on the learning rate, α (see Materials and Methods). These results suggest that observed levels of implicit biases emerge from the interaction between large-scale structural factors operating

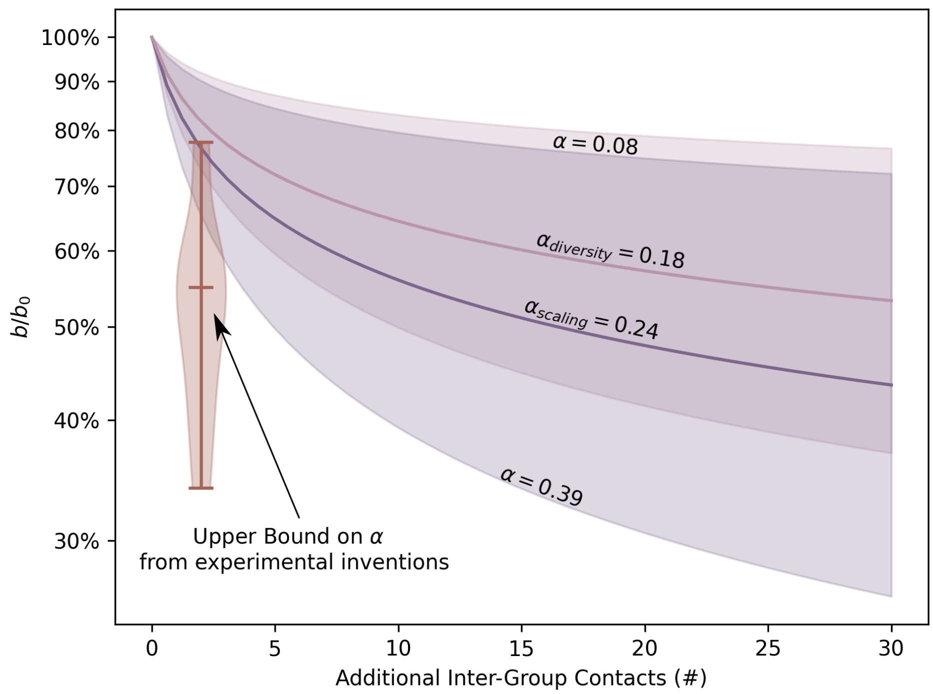


Figure 3: Estimated learning rates, α. We plot learning as a decrease in bias levels relative to an arbitrary baseline, $\frac{b}{b_0}$ as a function of the number of additional inter-group contacts. Solid curves indicate the mean estimated learning rate from the scaling exponent or majority group adjustment (diversity effect) averaged across years. Shaded regions show the 95% confidence intervals for the learning rate estimates with the lower envelop and upper envelope referring to the scaling exponent and diversity estimates, respectively. The violin plot gives an upper bound on the learning rate from 18 previously conducted experimental interventions (*25, 26*) designed to simulate one-shot inter-group contact of varying quality.

across entire cities to shape social contexts, and individual psychology which determines how much and how quickly people learn from and internalize those social contexts.

Discussion

The model developed here demonstrates that relatively simple considerations of heterogeneous mixing among a small number of social groups can explain a large proportion of why people in some cities have stronger implicit racial biases than in others. While it is somewhat surprising that only three factors - city population, majority group size, and racial segregation - account for so much between-city difference, this is in line with recent evidence that implicit racial biases are driven more by social contexts than by individual differences in attitudes (*50, 51*). Importantly, our model suggests that implicit racial biases emerge from the interaction between city-wide social contexts that are shaped by the built environment and individual psychology which determines how much and how quickly people learn from those contexts.

These effects suggest that as more people move into cities over the next decades, implicit biases will decrease, so long as cities do not become too segregated, remain centers of diversity, and residents continue to learn from shifting social environments. Though the amount of variance explained by segregation effects is small, reductions in implicit racial biases from decreasing segregation in cities could have large societal impacts (*52*). This is important to recognize as cities with lower levels of racial segregation also tend to, not accidentally, have higher incomes (*32*) and healthier inhabitants (*53*).

These results, along with our related work (*32*) characterizing economic productivity, are first steps towards better incorporating heterogeneous network structures and individual psychology into the mathematical models of modern urban science and deriving associated multifaceted effects. The additions we developed here are relatively simplistic in their consideration of individual differences in cities, proxied simply by a set of discrete groups. More complex models are likely needed to consider how city organization influences the dynamics of other types of attitudes that are socially relevant, including political polarization (*54, 55*) and issues of trust and collective action, for example relating to public health programs such as vaccines (*56, 57*).

References

1. L. M. Bettencourt, *Science* **340**, 1438 (2013).
2. L. M. Bettencourt, *Introduction to urban science: evidence and theory of cities as complex systems* (MIT Press, 2021).
3. C. Molinero, S. Thurner, *Journal of the Royal Society interface* **18**, 20200705 (2021).
4. M. Schläpfer, *et al.*, *Journal of the Royal Society Interface* **11**, 20130789 (2014).
5. M. Oliveira, C. Bastos-Filho, R. Menezes, *PloS one* **12**, e0183110 (2017).
6. J. Lobo, L. M. Bettencourt, M. E. Smith, S. Ortman, *Urban Studies* **57**, 731 (2020).
7. Y. Dunham, A. S. Baron, M. R. Banaji, *Child development* **77**, 1268 (2006).
8. E. Dehon, *et al.*, *Academic Emergency Medicine* **24**, 895 (2017).
9. D. S. Jacoby-Senghor, S. Sinclair, J. N. Shelton, *Journal of Experimental Social Psychology* **63**, 50 (2016).
10. J. C. Ziegert, P. J. Hanges, *Journal of applied psychology* **90**, 553 (2005).
11. P. D. Ekstrom, J. M. Le Forestier, C. K. Lai, *Psychological science* **33**, 497 (2022).
12. E. Hehman, J. K. Flake, J. Calanchini, *Social psychological and personality science* **9**, 393 (2018).

13. D. H. Chae, *et al.*, *Journal of Black Psychology* **43**, 789 (2017).

14. D. H. Chae, A. M. Nuru-Jeter, N. E. Adler, *Psychosomatic Medicine* **74**, 961 (2012).

15. A. S. Baron, M. R. Banaji, *Psychological science* **17**, 53 (2006).

16. B. L. Gibson, P. Rochat, E. B. Tone, A. S. Baron, *PloS one* **12**, e0183015 (2017).

17. J. T. Kubota, M. R. Banaji, E. A. Phelps, *Nature neuroscience* **15**, 940 (2012).

18. D. M. Amodio, *Nature Reviews Neuroscience* **15**, 670 (2014).

19. J. Cloutier, T. Li, B. Mišić, J. Correll, M. G. Berman, *Cerebral Cortex* **27**, 4326 (2017).

20. T. Brosch, E. Bar-David, E. A. Phelps, *Psychological science* **24**, 160 (2013).

21. B. K. Payne, H. A. Vuletich, *Policy Insights from the Behavioral and Brain Sciences* **5**, 49 (2018).

22. S. M. Jackson, A. L. Hillard, T. R. Schneider, *Social Psychology of Education* **17**, 419 (2014).

23. C. K. Boscardin, *Journal of General Internal Medicine* **30**, 1726 (2015).

24. A. M. Gonzalez, J. R. Steele, A. S. Baron, *Child Development* **88**, 123 (2017).

25. C. K. Lai, *et al.*, *Journal of Experimental Psychology: General* **143**, 1765 (2014).

26. C. K. Lai, *et al.*, *Journal of Experimental Psychology: General* **145**, 1001 (2016).

27. B. K. Payne, J. W. Hannay, *Trends in cognitive sciences* **25**, 927 (2021).

28. J. De Houwer, *Perspectives on Psychological Science* **14**, 835 (2019).

29. H. A. Vuletich, B. K. Payne, *Psychological science* **30**, 854 (2019).

30. B. K. Payne, H. A. Vuletich, J. L. Brown-Iannuzzi, *Proceedings of the National Academy of Sciences* **116**, 11693 (2019).

31. L. M. Bettencourt, J. Lobo, D. Strumsky, G. B. West, *PloS one* **5**, e13541 (2010).

32. A. J. Stier, S. Sajjadi, L. M. A. Bettencourt, F. Karimi, M. G. Berman, *arXiv* (2022).

33. J. Laurence, *European Sociological Review* **27**, 70 (2011).

34. T. F. Pettigrew, O. Christ, U. Wagner, J. Stellmacher, *International Journal of intercultural relations* **31**, 411 (2007).

35. G. W. Allport, K. Clark, T. Pettigrew, *The nature of prejudice* (Addison-wesley Reading, MA, 1954).

36. U. Wagner, R. Van Dick, T. F. Pettigrew, O. Christ, *Group Processes & Intergroup Relations* **6**, 22 (2003).

37. U. Wagner, O. Christ, T. F. Pettigrew, J. Stellmacher, C. Wolf, *Social psychology quarterly* **69**, 380 (2006).

38. E. R. Crossman, *Ergonomics* **2**, 153 (1959).

39. P. A. Woźniak, E. J. Gorzelańczyk, J. A. Murakowski, *Acta neurobiologiae experimentalis* **55**, 301 (1995).

40. J. M. Murre, J. Dros, *PloS one* **10**, e0120644 (2015).

41. B. Van der Zwaan, A. Rabl, *Solar energy* **74**, 19 (2003).

42. R. Cesaretti, J. Lobo, L. M. Bettencourt, S. G. Ortman, M. E. Smith, *PloS one* **11**, e0162678 (2016).

43. L. M. Bettencourt, *Science advances* **6**, eaat8812 (2020).

44. K. Xu, B. Nosek, A. Greenwald, *Journal of Open Psychology Data* **2** (2014).

45. A. J. Stier, *et al.*, *Proceedings of the National Academy of Sciences* **119**, e2119313118 (2022).

46. E. Hehman, E. K. Ofosu, J. Calanchini, *Social Psychological and Personality Science* **12**, 156 (2021).

47. R. Tucker, *et al.*, *Journal of Quantitative Criminology* **37**, 333 (2021).

48. K. R. Storrs, S.-M. Khaligh-Razavi, N. Kriegeskorte, *BioRxiv* (2020).

49. K. Schnabel, J. B. Asendorpf, A. G. Greenwald, *European Journal of Psychological Assessment* **24**, 210 (2008).

50. B. K. Payne, H. A. Vuletich, K. B. Lundberg, *Psychological Inquiry* **28**, 233 (2017).

51. E. Lee, *et al.*, *Nature human behaviour* **3**, 1078 (2019).

52. A. G. Greenwald, M. R. Banaji, B. A. Nosek, *Journal of Personality and Social Pyschology* (2015).

53. M. R. Kramer, C. R. Hogue, *Epidemiologic reviews* **31**, 178 (2009).

54. J. B. Bak-Coleman, *et al.*, *Proceedings of the National Academy of Sciences* **118**, e2025764118 (2021).

55. J. Dalege, D. Borsboom, F. van Harreveld, L. J. Waldorp, H. L. van der Maas, *Scientific reports* **7**, 1 (2017).

56. M. Galesic, H. Olsson, J. Dalege, T. van der Does, D. L. Stein, *Journal of the Royal Society Interface* **18**, 20200857 (2021).

57. J. Dalege, T. van der Does, *Science Advances* **8**, eabm0137 (2022).

58. A. G. Greenwald, B. A. Nosek, M. R. Banaji, *Journal of personality and social psychology* **85**, 197 (2003).

59. M. J. White, *Population index* pp. 198–221 (1986).

60. A. J. Kind, W. R. Buckingham, *The New England journal of medicine* **378**, 2456 (2018).

61. U. of Wisconsin School of Medicine, P. Health., Area deprivation index 2019.

62. A. Lage-Castellanos, G. Valente, E. Formisano, F. De Martino, *PLoS computational biology* **15**, e1006397 (2019).

Acknowledgments

The authors declare that they have no competing interests. All data needed to evaluate the conclusions in the paper are present in the paper, the Supplementary Materials, or are publicly available.

Author contributions: A.J.S., L.M.A.B and M.G.B designed research; A.J.S. performed research; M.G.B supervised research; and A.J.S., S.S, F.K., L.M.A.B, and M.G.B. wrote the paper.

This work is partially supported by NSF-2106013, and S&CC-1952050.

Materials and Methods

IAT Data

All racial IAT Data are publicly available (*44*) and were downloaded from `https://osf.io/52qxl/`. These data are coded at the participant level, a fraction of which include geographic identifiers for state and county. Implicit racial bias was assessed by the D_{biep} metric (*58*) which consists of the latency difference between compatible and incompatible blocks of the racial IAT, divided by the pooled standard deviation. In the racial IAT, black and white face images are used and higher and positive D_{biep} scores indicate an implicit bias towards white faces while lower and negative D_{biep} scores indicate an implicit bias towards black faces. After only retaining participants with available geographic information, D_{biep} scores were averaged across all participants in each CBSA. Cities were retained if they had at least 500 IAT responses. This was done separately for all years. Results were similar with cutoffs of > 250 and > 1000 IAT responses per city (Supplementary Tables 37-44).

U.S. Census Data

All census data is publicly available and was downloaded from `data.census.gov`. Five-year racial demographic estimates for U.S. census tracts were downloaded from table *B02001*. Heterophobia values were calculated across the two racial groups in the race IAT: White and Black. Five-year population estimates for U.S. cities defined as combined statistical areas (CBSAs) were downloaded from table *B01003*. In order to map between census tracts and CBSAs, delineation files for 2020 were downloaded from the United States Office of Budget and Management from `https://www.census.gov/programs-surveys/metro-micro/about/delineation-files.html`.

Associations Between Implicit Bias, City Size, Majority Group Size, and Heterophobia

We fit the scaling law between the logarithms of implicit bias and city size with ordinary least squares (OLS) linear regression to determine the scaling exponent. The equation for this regression is:

$$\ln(b_i) \sim C + \beta_1 \cdot \ln(N_i) + \epsilon_i \tag{4}$$

where C is the log-log intercept (or equivalently the logarithm of the scaling prefactor), β_1 log-log slope (i.e., the scaling exponent), and ϵ_i are the scaling deviations.

In order to assess the contribution of the city-specific majority group size and heterophobia values to implicit racial bias, we start with ϵ_i as the dependent variable via the equation:

$$\epsilon_i \sim C_2 + \beta_2 \cdot \ln(\frac{N_{1,i}}{N_i} - \frac{N_{1,i}^2}{N_i^2}) + \beta_3 \cdot \ln(2 - h_{1,i} - h_{2,i}) + \xi_i \tag{5}$$

where $N_{1,i}$ is the number of white individuals city i, $h_{1,i}$ is the heterophobia of the white population, and $h_{2,i}$ is the heterophobia of the black population, and ξ_i are additional city specific residual effects.

Since we do not observe heterophobia values, $h_{1,i}$ and $h_{2,i}$, directly, but only measures of residential racial segregation, $s_{1,i}$ and $s_{2,i}$, we follow our related work (*32*) and model the heterophobia values as linearly dependent on levels of residential racial segregation. With the additional approximation that $\ln(2-x) \simeq \ln(2) - \frac{x}{2}$ when $x << 1$, equation 5 then becomes:

$$\epsilon_i \simeq C_2 + \beta_2 \cdot \ln(\frac{N_{1,i}}{N_i} - \frac{N_{1,i}^2}{N_i^2}) - \frac{\beta_3}{2} \cdot [2 \cdot h^{het} + b^{het} \cdot (s_{1,i} + s_{2,i})] + \beta_3 \ln(2) + \xi_i \tag{6}$$

where we have substituted the heterophobia values via the equation $h_{g,i} = h^{het} + b^{het} s_{g,i}$ (*32*). We can further simplify by including all non-city specific effects in the constant C_2 and by including the factor of $\frac{-b^{het}}{2}$ in the constant, β_3. We fit the resulting equation via OLS in order to assess the contribution of majority group size and residential racial segregation to implicit racial bias:

$$\epsilon_i \simeq C_2 + \beta_2 \cdot \ln(\frac{N_{1,i}}{N_i} - \frac{N_{1,i}^2}{N_i^2}) + \beta_3 \cdot (s_{1,i} + s_{2,i}) + \xi_i \quad (7)$$

Noise Ceiling Estimates

In order to estimate the noise ceiling, we computed the correlation between IAT bias measures between halves for 500 split permutations of individual IAT participants in each year. The upper bound of the noise ceiling was estimated by averaging the correlations between each half and the full sample, while the lower bound of the noise ceiling was estimated by correlating IAT bias between the two halves of each split half (*48*).

Measures of Residential Segregation

As in our related work (*32*), all analyses were conducted across four different measures of residential segregation (*59*) in order to ensure that the results were not sensitive to any specific metric. These included the mean deviance measure:

$$\Delta_{g,i} = \frac{1}{M} \sum_{m}^{M} |N_{g,m,i}/N_{m,i} - N_{g,i}/N_i|, \quad (8)$$

the normalized segregation index:

$$D_{g,i} = \frac{\sum_m |\frac{N_{g,m,i}}{N_{m,i}} - \frac{N_{g,i}}{N_i}| \cdot N_{m,i}}{2 \cdot N_i \cdot \frac{N_{g,i}}{N_i} \cdot (1 - \frac{N_{g,i}}{N_i})}, \quad (9)$$

the Gini Coefficient:

$$gini_{g,i} = \frac{\sum_m \sum_l |\frac{N_{g,m,i}}{N_{m,i}} - \frac{N_{g,l,i}}{N_{l,i}}| \cdot N_{m,i} \cdot N_{l,i}}{2 \cdot N_i^2 \cdot \frac{N_{g,i}}{N_i} \cdot (1 - \frac{N_{g,i}}{N_i})}, \quad (10)$$

and the exposure B_{gg} index, also known as the correlation ratio (CR or η^2) or the mean squared deviation:

$$\eta_{g,i}^2 = \frac{\sum_m N_{g,m,i}^2}{N_{g,i} \cdot (1 - \frac{N_{g,i}}{N_i})} - \frac{\frac{N_{g,i}}{N_i}}{1 - \frac{N_{g,i}}{N_i}}. \quad (11)$$

Controlling For Individual Demographics

In order to control for individual demographics of IAT respondents, we transformed the individual bias responses into an indicator for $D_{biep} > 0$. This variable thus indicates whether the individual respondent had a positive bias for white faces or not. For each year, a logistic regression was performed that included the city-level variables of the natural logarithm of population, the majority groups size adjustment, and the heterophobia adjustment, and the individual level variables of race, educational attainment and birth sex. The 14-point educational attainment scale included with the IAT data, *edu_14*, was recoded into three categories of "High School Graduate or Below", "Some College or College Graduate", and "Advanced Degree". For some years there were no respondents in the "High School Graduate or Below" category, in which case that variable was excluded from analyses. Self reported racial demographics (*raceomb* before 2016 and *raceomb_002* afterwards) was recoded to three categories of "White", "Black", and "Multiracial", with other races and "unknown" combined as the base category.

Comparison to Previous Results Associating Area Deprivation With Racial IAT Responses

We downloaded the average maximum heat index (HI) in degrees Celsius for U.S. counties from the North America Land Data Assimilation System Daily Air Temperatures and Heat Index 1979-2011 database. This was the strongest predictor of between-city differences in implicit racial bias levels in a previously published analysis (*46*). The maximum heat index was averaged across counties within each CBSA.

Those analyses used a "kitchen-sink" approach with regularizing regressions to determine which variables were relevant to predicting these differences between cities. Since the variables identified there are indicative of levels of environmental, social, and economic disadvantage, we additionally evaluated the relevance of the Area Deprivation Index (ADI) to between-city differences in implicit racial bias. The ADI measures neighborhood socioeconomic disadvantage at small spatial units down to the census block level and includes factors related to income, education, employment, and housing quality (*60, 61*). We averaged nationally anchored ADI values at the county level across all counties in each CBSA.

In order to determine the effects of these measures of neighborhood disadvantage on implicit racial biases we conducted separate OLS regressions including city size, the majority group size adjustment, the heterophobia adjustment, and the ADI or HI. Since ADI and HI data are not available for all CBSAs, we additionally conducted regressions without the ADI and HI included, but with the reduced sample size for which these data are available. We note that in those regressions with a reduced sample size, but without the inclusion of the ADI or HI the variance explained by city size, the majority group size adjustment, the heterophobia adjustment are higher than in the full sample, and outperform previous analyses which only include measures of neighborhood disadvantage (*46*).

Estimates of the Learning Rate

Independent empirical estimates of the learning rate, α, which governs the coupling between inter-group interactions and bias levels, were obtained directly from the two-step OLS regressions described in Equations 5 and 7. From Equation 5 we obtain an estimate of $\hat{\alpha}_{scaling} = \frac{\beta_1}{\delta}$. Confidence intervals for $\hat{\alpha}_{scaling}$ were obtained from the OLS confidence intervals for β_1. We note there may be other effects besides learning such as top-down hierarchical structures and variations in growth rates that may additionally contribute to differences in the empirical scaling exponent β_1 from the expected value of $\delta = \frac{1}{6}$. In addition, we obtain a second, independent estimate of the learning rate: $\hat{\alpha}_{diversity} = \beta_2$ based on Equation 3 of the main text.

Results from experimental interventions designed to simulate inter-group contact were used to further validate and bound these estimates of α. We calculated the relative reduction in

IAT D_{biep} scores pre- and post-intervention for 18 different systematic interventions of various strength (*25, 62*). These interventions included having participants read stories of various lengths and vividness designed to affirm white-bad and black-good associations, modifying the IAT to include additional black-good and white-bad blocks, simulating competition with white opponents and cooperation with black teammates, having participants read about threatening scenarios and shown images of "friends" in those scenarios and reminding participants of prominent black athletes positives contributions to society (*25, 26*). Importantly, all of these interventions occurred directly between IAT tests and are all positive in nature. In reality, inter-group interactions may not always be positive in nature, and they play out continuously at potentially irregular intervals relative to when a given individual makes a judgment or decision that is influenced by implicit racial biases. Consequently, these experimental interventions can be interpreted as an upper bound on the effects of one additional inter-group interaction when that interaction happens immediately before implicit bias levels are assessed.

Supplementary Text

Inter-group interactions with two groups

In the case of two groups, the inter-group term of Equation 1 of the main text becomes:

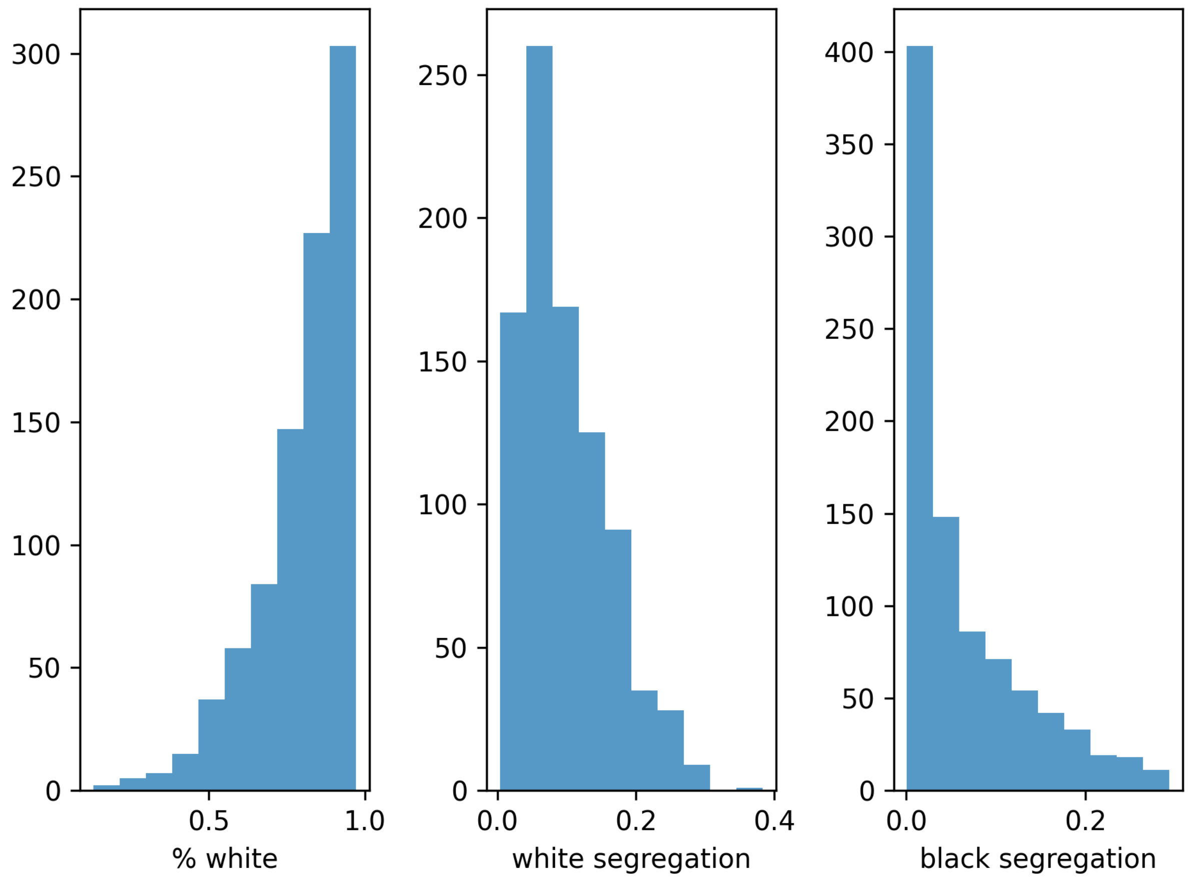
$$k_{inter,i} \sim N_i^{\delta}[2 \cdot \frac{N_{1,i}}{N_i}\frac{N_{2,i}}{N_i} - \frac{N_{1,i}}{N_i}\frac{N_{2,i}}{N_i}h_{1,i}^{het} - \frac{N_{1,i}}{N_i}\frac{N_{2,i}}{N_i}h_{2,i}^{het}] \tag{1}$$

Replacing $N_{2,i}$ with $N_i - N_{1,i}$ and simplifying results in:

$$k_i \sim N^{\delta}[(\frac{N_{1,i}}{N_i}\frac{N_i - N_{1,i}}{N_i}) \cdot (2 - h_{1,i}^{het} - h_{2,i}^{het})] \tag{2}$$

which simplifies directly to Equation 3 of the main text when using a power-law learning curve to couple inter-croup interactions to implicit bias levels.

Supplementary Figures



Supplementary Figure 1: Histograms of percent white, white residential racial segregation, and black residential racial segregation for CBSAs in 2020. The y-axis denotes the number of cities in a given histogram bin.

Supplementary Tables

Supplementary Table 1: Summary of scaling fits and majority group size adjustment and heterophobia adjustment parameters for cities with more than 500 IAT responses.

year	scaling β_1	majority group size adjustment β_2	$\mathbf{s_1} + \mathbf{s_2}\ \beta_3$	# cities
2010	[-0.094,-0.021]	[-0.277,0.038]	[-0.563,0.128]	64
2011	[-0.086,-0.010]	[-0.296,0.048]	[-0.526,0.201]	59
2012	[-0.080,-0.013]	[-0.334,-0.078]	[-0.200,0.315]	46
2013	[-0.076,-0.019]	[-0.327,-0.064]	[-0.174,0.361]	52
2014	[-0.069,-0.016]	[-0.221,0.027]	[-0.321,0.225]	66
2015	[-0.047,-0.003]	[-0.305,-0.130]	[0.054,0.447]	74
2016	[-0.055,-0.015]	[-0.264,-0.095]	[0.032,0.415]	91
2017	[-0.059,-0.016]	[-0.308,-0.128]	[0.055,0.498]	105
2018	[-0.061,-0.019]	[-0.348,-0.158]	[0.105,0.563]	102
2019	[-0.063,-0.017]	[-0.292,-0.094]	[-0.073,0.409]	106
2020	[-0.044,-0.011]	[-0.278,-0.133]	[0.105,0.466]	149
all years	[-0.045,-0.031]	[-0.226,-0.163]	[0.026,0.066]	914

Supplementary Table 2: IAT participants with geographic information

year	IAT sample size	median CBSA population %
2010	163,289	0.09%
2011	149,059	0.08%
2012	118,179	0.06%
2013	125,083	0.07%
2014	171,039	0.09%
2015	208,521	0.11%
2016	273,995	0.14%
2017	321,082	0.17%
2018	316,393	0.17%
2019	321,247	0.16%
2020	584,902	0.26%

Supplementary Table 3: Logistic Regression to predict individual racial IAT bias scores > 0 for 2010. Note that larger and less segregated cities are associated with a lower probability of a positive bias towards white faces, in line with Equation 3 of the main text. More diversity, captured by majority group size is significant here.

Dep. Variable:	bias	**No. Observations:**	108303
Model:	Logit	**Df Residuals:**	108293
Method:	MLE	**Df Model:**	9
		Pseudo R-squ.:	0.08710
		Log-Likelihood:	-53918.
converged:	True	**LL-Null:**	-59062.
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	0.9954	5.67e+05	1.75e-06	1.000	-1.11e+06	1.11e+06
ln(population)	-0.0195	0.009	-2.057	0.040	-0.038	-0.001
White	0.4831	0.024	20.211	0.000	0.436	0.530
Black	-1.3555	0.026	-51.153	0.000	-1.407	-1.304
Multiracial	-0.1467	0.039	-3.740	0.000	-0.224	-0.070
High School or Less	0.2787	5.67e+05	4.91e-07	1.000	-1.11e+06	1.11e+06
College	0.3061	5.67e+05	5.4e-07	1.000	-1.11e+06	1.11e+06
Advanced Degree	0.4106	5.67e+05	7.24e-07	1.000	-1.11e+06	1.11e+06
Majority Group Size Adjustment	0.0115	0.049	0.234	0.815	-0.085	0.108
$\mathbf{s_1 + s_2}$	0.2599	0.084	3.077	0.002	0.094	0.426

Supplementary Table 4: Logistic Regression to predict individual racial IAT bias scores > 0 for 2011. Note that larger and less segregated cities are associated with a lower probability of a positive bias towards white faces, in line with Equation 3 of the main text. More diversity, captured by majority group size is significant here.

Dep. Variable:	bias	**No. Observations:**	207358
Model:	Logit	**Df Residuals:**	207348
Method:	MLE	**Df Model:**	9
		Pseudo R-squ.:	0.08369
		Log-Likelihood:	-1.0348e+05
converged:	True	**LL-Null:**	-1.1293e+05
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P$>$ $\lvert$z$\rvert$**	**[0.025**	**0.975]**
const	1.0078	1.58e+05	6.37e-06	1.000	-3.1e+05	3.1e+05
ln(population)	-0.0172	0.007	-2.525	0.012	-0.031	-0.004
White	0.4839	0.017	28.343	0.000	0.450	0.517
Black	-1.3244	0.019	-69.748	0.000	-1.362	-1.287
Multiracial	-0.1428	0.028	-5.085	0.000	-0.198	-0.088
High School or Less	0.2770	1.58e+05	1.75e-06	1.000	-3.1e+05	3.1e+05
College	0.3089	1.58e+05	1.95e-06	1.000	-3.1e+05	3.1e+05
Advanced Degree	0.4219	1.58e+05	2.67e-06	1.000	-3.1e+05	3.1e+05
Majority Group Size Adjustment	0.0347	0.036	0.975	0.329	-0.035	0.104
$\mathbf{s_1 + s_2}$	0.1951	0.061	3.200	0.001	0.076	0.315

Supplementary Table 5: Logistic Regression to predict individual racial IAT bias scores > 0 for 2012. Note that larger and less segregated cities are associated with a lower probability of a positive bias towards white faces, in line with Equation 3 of the main text. More diversity, captured by majority group size is significant here.

Dep. Variable:	bias	**No. Observations:**	281203
Model:	Logit	**Df Residuals:**	281194
Method:	MLE	**Df Model:**	8
		Pseudo R-squ.:	0.08122
		Log-Likelihood:	-1.4032e+05
converged:	True	**LL-Null:**	-1.5272e+05
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P$>\|z\|$**	**[0.025**	**0.975]**
const	1.1506	0.131	8.811	0.000	0.895	1.407
ln(population)	-0.0125	0.006	-2.083	0.037	-0.024	-0.001
White	0.4821	0.015	33.173	0.000	0.454	0.511
Black	-1.3093	0.016	-80.596	0.000	-1.341	-1.277
Multiracial	-0.1456	0.024	-6.057	0.000	-0.193	-0.098
College	0.0699	0.036	1.939	0.053	-0.001	0.141
Advanced Degree	0.1605	0.023	6.971	0.000	0.115	0.206
Majority Group Size Adjustment	0.0095	0.031	0.307	0.759	-0.051	0.070
$\mathbf{s_1 + s_2}$	0.2116	0.052	4.039	0.000	0.109	0.314

Supplementary Table 6: Logistic Regression to predict individual racial IAT bias scores > 0 for 2013. Note that larger and less segregated cities are associated with a lower probability of a positive bias towards white faces, in line with Equation 3 of the main text. More diversity, captured by majority group size is significant here.

Dep. Variable:	bias	**No. Observations:**	363764
Model:	Logit	**Df Residuals:**	363754
Method:	MLE	**Df Model:**	9
		Pseudo R-squ.:	0.07789
		Log-Likelihood:	-1.8202e+05
converged:	True	**LL-Null:**	-1.9740e+05
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P$>|z|$**	**[0.025**	**0.975]**
const	0.9616	3.98e+05	2.41e-06	1.000	-7.8e+05	7.8e+05
ln(population)	-0.0160	0.005	-3.013	0.003	-0.026	-0.006
White	0.4680	0.013	37.078	0.000	0.443	0.493
Black	-1.2990	0.014	-91.452	0.000	-1.327	-1.271
Multiracial	-0.1638	0.021	-7.817	0.000	-0.205	-0.123
High School or Less	0.2269	3.98e+05	5.7e-07	1.000	-7.8e+05	7.8e+05
College	0.3283	3.98e+05	8.25e-07	1.000	-7.8e+05	7.8e+05
Advanced Degree	0.4063	3.98e+05	1.02e-06	1.000	-7.8e+05	7.8e+05
Majority Group Size Adjustment	0.0125	0.028	0.450	0.653	-0.042	0.067
$\mathbf{s_1 + s_2}$	0.2257	0.046	4.918	0.000	0.136	0.316

Supplementary Table 7: Logistic Regression to predict individual racial IAT bias scores > 0 for 2014. Note that larger and less segregated cities are associated with a lower probability of a positive bias towards white faces, in line with Equation 3 of the main text. More diversity, captured by majority group size is significant here.

Dep. Variable:	bias	**No. Observations:**	487488
Model:	Logit	**Df Residuals:**	487478
Method:	MLE	**Df Model:**	9
		Pseudo R-squ.:	0.07078
		Log-Likelihood:	-2.4760e+05
converged:	True	**LL-Null:**	-2.6645e+05
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P$>\|z\|$**	**[0.025**	**0.975]**
const	0.8539	1.11e+05	7.69e-06	1.000	-2.18e+05	2.18e+05
ln(population)	-0.0144	0.004	-3.210	0.001	-0.023	-0.006
White	0.4435	0.011	41.192	0.000	0.422	0.465
Black	-1.2695	0.012	-103.076	0.000	-1.294	-1.245
Multiracial	-0.1576	0.018	-8.734	0.000	-0.193	-0.122
High School or Less	0.1917	1.11e+05	1.73e-06	1.000	-2.18e+05	2.18e+05
College	0.2784	1.11e+05	2.51e-06	1.000	-2.18e+05	2.18e+05
Advanced Degree	0.3838	1.11e+05	3.46e-06	1.000	-2.18e+05	2.18e+05
Majority Group Size Adjustment	-0.0242	0.023	-1.033	0.301	-0.070	0.022
$\mathbf{s_1 + s_2}$	0.2678	0.039	6.817	0.000	0.191	0.345

Supplementary Table 8: Logistic Regression to predict individual racial IAT bias scores > 0 for 2015. Note that larger and less segregated cities are associated with a lower probability of a positive bias towards white faces, in line with Equation 3 of the main text. More diverse cities (captured by majority group size) cities are trending in the direction of less bias, but are not significant here.

Dep. Variable:	bias	**No. Observations:**	518942
Model:	Logit	**Df Residuals:**	518933
Method:	MLE	**Df Model:**	8
		Pseudo R-squ.:	0.06926
		Log-Likelihood:	-2.6465e+05
converged:	True	**LL-Null:**	-2.8435e+05
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	0.9648	0.094	10.305	0.000	0.781	1.148
ln(population)	-0.0125	0.004	-2.888	0.004	-0.021	-0.004
White	0.4411	0.010	42.375	0.000	0.421	0.461
Black	-1.2592	0.012	-105.380	0.000	-1.283	-1.236
Multiracial	-0.1573	0.017	-9.010	0.000	-0.192	-0.123
College	0.0921	0.025	3.711	0.000	0.043	0.141
Advanced Degree	0.1949	0.015	13.191	0.000	0.166	0.224
Majority Group Size Adjustment	-0.0425	0.023	-1.886	0.059	-0.087	0.002
$\mathbf{s_1 + s_2}$	0.2889	0.038	7.630	0.000	0.215	0.363

Supplementary Table 9: Logistic Regression to predict individual racial IAT bias scores > 0 for 2016. Note that larger, more diverse, and less segregated cities are associated with a lower probability of a positive bias towards white faces, in line with Equation 3 of the main text.

Dep. Variable:	bias	**No. Observations:**	79309
Model:	Logit	**Df Residuals:**	79299
Method:	MLE	**Df Model:**	9
		Pseudo R-squ.:	0.05319
		Log-Likelihood:	-40512.
converged:	True	**LL-Null:**	-42787.
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P$>|z|$**	**[0.025**	**0.975]**
const	1.0349	0.203	5.106	0.000	0.638	1.432
ln(population)	-0.0404	0.010	-4.129	0.000	-0.060	-0.021
White	0.4224	0.028	15.098	0.000	0.368	0.477
Black	-1.0940	0.033	-33.631	0.000	-1.158	-1.030
Multiracial	-0.3113	0.042	-7.375	0.000	-0.394	-0.229
Birth Sex	0.1161	0.018	6.412	0.000	0.081	0.152
College	0.2359	0.053	4.440	0.000	0.132	0.340
Advanced Degree	0.2345	0.033	7.032	0.000	0.169	0.300
Majority Group Size Adjustment	-0.1155	0.053	-2.181	0.029	-0.219	-0.012
$\mathbf{s_1 + s_2}$	0.7081	0.099	7.148	0.000	0.514	0.902

Supplementary Table 10: Logistic Regression to predict individual racial IAT bias scores > 0 for 2017. Note that larger, more diverse, and less segregated cities are associated with a lower probability of a positive bias towards white faces, in line with Equation 3 of the main text.

Dep. Variable:	bias	**No. Observations:**	328522
Model:	Logit	**Df Residuals:**	328512
Method:	MLE	**Df Model:**	9
		Pseudo R-squ.:	0.05032
		Log-Likelihood:	-1.7217e+05
converged:	True	**LL-Null:**	-1.8129e+05
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	0.7606	0.102	7.431	0.000	0.560	0.961
ln(population)	-0.0204	0.005	-4.214	0.000	-0.030	-0.011
White	0.3824	0.013	28.393	0.000	0.356	0.409
Black	-1.1044	0.016	-69.596	0.000	-1.135	-1.073
Multiracial	-0.3722	0.020	-18.657	0.000	-0.411	-0.333
Birth Sex	0.1480	0.009	16.815	0.000	0.131	0.165
College	0.1414	0.026	5.489	0.000	0.091	0.192
Advanced Degree	0.1389	0.017	8.226	0.000	0.106	0.172
Majority Group Size Adjustment	-0.1516	0.026	-5.926	0.000	-0.202	-0.101
$\mathbf{s_1 + s_2}$	0.5886	0.047	12.448	0.000	0.496	0.681

Supplementary Table 11: Logistic Regression to predict individual racial IAT bias scores > 0 for 2018. Note that larger, more diverse, and less segregated cities are associated with a lower probability of a positive bias towards white faces, in line with Equation 3 of the main text.

Dep. Variable:	bias	**No. Observations:**	570868
Model:	Logit	**Df Residuals:**	570858
Method:	MLE	**Df Model:**	9
		Pseudo R-squ.:	0.05081
		Log-Likelihood:	-3.0335e+05
converged:	True	**LL-Null:**	-3.1959e+05
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P$>|z|$**	**[0.025**	**0.975]**
const	0.5022	0.078	6.444	0.000	0.349	0.655
ln(population)	-0.0124	0.004	-3.384	0.001	-0.020	-0.005
White	0.4082	0.010	40.840	0.000	0.389	0.428
Black	-1.0764	0.012	-91.337	0.000	-1.100	-1.053
Multiracial	-0.3381	0.015	-22.647	0.000	-0.367	-0.309
Birth Sex	0.1426	0.007	21.479	0.000	0.130	0.156
College	0.1629	0.019	8.368	0.000	0.125	0.201
Advanced Degree	0.1423	0.013	11.026	0.000	0.117	0.168
Majority Group Size Adjustment	-0.1937	0.019	-9.958	0.000	-0.232	-0.156
$\mathbf{s_1 + s_2}$	0.6276	0.036	17.633	0.000	0.558	0.697

Supplementary Table 12: Logistic Regression to predict individual racial IAT bias scores > 0 for 2019. Note that larger, more diverse, and less segregated cities are associated with a lower probability of a positive bias towards white faces, in line with Equation 3 of the main text.

Dep. Variable:	bias	**No. Observations:**	570868
Model:	Logit	**Df Residuals:**	570858
Method:	MLE	**Df Model:**	9
		Pseudo R-squ.:	0.05081
		Log-Likelihood:	-3.0335e+05
converged:	True	**LL-Null:**	-3.1959e+05
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	0.5022	0.078	6.444	0.000	0.349	0.655
ln(population)	-0.0124	0.004	-3.384	0.001	-0.020	-0.005
White	0.4082	0.010	40.840	0.000	0.389	0.428
Black	-1.0764	0.012	-91.337	0.000	-1.100	-1.053
Multiracial	-0.3381	0.015	-22.647	0.000	-0.367	-0.309
Birth Sex	0.1426	0.007	21.479	0.000	0.130	0.156
College	0.1629	0.019	8.368	0.000	0.125	0.201
Advanced Degree	0.1423	0.013	11.026	0.000	0.117	0.168
Majority Group Size Adjustment	-0.1937	0.019	-9.958	0.000	-0.232	-0.156
$\mathbf{s_1 + s_2}$	0.6276	0.036	17.633	0.000	0.558	0.697

Supplementary Table 13: Logistic Regression to predict individual racial IAT bias scores > 0 for 2020. Note that larger, more diverse, and less segregated cities are associated with a lower probability of a positive bias towards white faces, in line with Equation 3 of the main text.

Dep. Variable:	bias	**No. Observations:**	570868
Model:	Logit	**Df Residuals:**	570858
Method:	MLE	**Df Model:**	9
		Pseudo R-squ.:	0.05081
		Log-Likelihood:	-3.0335e+05
converged:	True	**LL-Null:**	-3.1959e+05
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	0.5022	0.078	6.444	0.000	0.349	0.655
ln(population)	-0.0124	0.004	-3.384	0.001	-0.020	-0.005
White	0.4082	0.010	40.840	0.000	0.389	0.428
Black	-1.0764	0.012	-91.337	0.000	-1.100	-1.053
Multiracial	-0.3381	0.015	-22.647	0.000	-0.367	-0.309
Birth Sex	0.1426	0.007	21.479	0.000	0.130	0.156
College	0.1629	0.019	8.368	0.000	0.125	0.201
Advanced Degree	0.1423	0.013	11.026	0.000	0.117	0.168
Majority Group Size Adjustment	-0.1937	0.019	-9.958	0.000	-0.232	-0.156
$\mathbf{s_1 + s_2}$	0.6276	0.036	17.633	0.000	0.558	0.697

Supplementary Table 14: Comparison of Models for cities that have available Area Deprivation Index (ADI) and Heat Index (HI) data. All models include city size, majority group size and heterophobia effects (mean deviation segregation measure).

year	no ADI R^2	ADI R^2	ADI n	no HI R^2	HI R^2	HI n
2010	0.383	0.394	36	0.329	0.357	18
2011	0.290	0.291	34	0.258	0.265	17
2012	0.370	0.389	27	0.463	0.478	17
2013	0.346	0.347	32	0.417	0.418	17
2014	0.248	0.259	39	0.368	0.448	18
2015	0.212	0.215	42	0.518	0.579	18
2016	0.150	0.162	50	0.502	0.546	19
2017	0.145	0.145	54	0.381	0.381	20
2018	0.204	0.222	53	0.468	0.496	19
2019	0.202	0.215	58	0.416	0.431	20
2020	0.160	0.161	76	0.437	0.438	22

Supplementary Table 15: Comparison of Models for cities that have available Area Deprivation Index (ADI) and Heat Index (HI) data. All models include city size, majority group size and heterophobia effects (segregation index).

year	no ADI R^2	ADI R^2	ADI n	no HI R^2	HI R^2	HI n
2010	0.399	0.412	36	0.376	0.397	18
2011	0.318	0.318	34	0.256	0.262	17
2012	0.418	0.438	27	0.517	0.527	17
2013	0.414	0.414	32	0.508	0.508	17
2014	0.289	0.300	39	0.445	0.516	18
2015	0.332	0.334	42	0.585	0.616	18
2016	0.213	0.231	50	0.639	0.662	19
2017	0.193	0.193	54	0.559	0.562	20
2018	0.273	0.293	53	0.671	0.690	19
2019	0.293	0.310	58	0.656	0.664	20
2020	0.288	0.288	76	0.650	0.650	22

Supplementary Table 16: Comparison of Models for cities that have available Area Deprivation Index (ADI) and Heat Index (HI) data. All models include city size, majority group size and heterophobia effects (gini coefficient).

year	no ADI R^2	ADI R^2	ADI n	no HI R^2	HI R^2	HI n
2010	0.401	0.416	36	0.366	0.388	18
2011	0.327	0.327	34	0.257	0.263	17
2012	0.424	0.448	27	0.518	0.529	17
2013	0.414	0.415	32	0.502	0.502	17
2014	0.291	0.300	39	0.442	0.517	18
2015	0.333	0.338	42	0.581	0.617	18
2016	0.223	0.239	50	0.633	0.660	19
2017	0.197	0.197	54	0.568	0.569	20
2018	0.271	0.288	53	0.665	0.686	19
2019	0.296	0.310	58	0.648	0.658	20
2020	0.296	0.296	76	0.651	0.652	22

Supplementary Table 17: Comparison of Models for cities that have available Area Deprivation Index (ADI) and Heat Index (HI) data. All models include city size, majority group size and heterophobia effects (η^2).

year	no ADI R^2	ADI R^2	ADI n	no HI R^2	HI R^2	HI n
2010	0.394	0.406	36	0.365	0.387	18
2011	0.310	0.310	34	0.252	0.259	17
2012	0.410	0.429	27	0.505	0.515	17
2013	0.406	0.406	32	0.490	0.490	17
2014	0.289	0.300	39	0.426	0.498	18
2015	0.293	0.296	42	0.589	0.623	18
2016	0.187	0.202	50	0.582	0.609	19
2017	0.178	0.179	54	0.494	0.497	20
2018	0.246	0.265	53	0.575	0.590	19
2019	0.258	0.270	58	0.556	0.561	20
2020	0.225	0.225	76	0.561	0.563	22

Supplementary Table 18: Summary of scaling fits and majority group size adjustment and heterophobia variance explained for cities with more than 500 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.122	0.180	0.008	0.327	64
2011	0.086	0.134	-0.003	0.246	59
2012	0.130	0.266	-0.014	0.400	46
2013	0.164	0.194	-0.008	0.368	52
2014	0.125	0.105	-0.012	0.242	66
2015	0.053	0.216	0.055	0.336	74
2016	0.112	0.121	0.041	0.277	91
2017	0.097	0.155	0.040	0.295	105
2018	0.114	0.170	0.057	0.335	102
2019	0.094	0.147	0.007	0.257	106
2020	0.061	0.128	0.049	0.242	149
all years	0.105	0.160	0.008	0.267	914

Supplementary Table 19: Summary of scaling fits and majority group size adjustment and heterophobia variance explained estimated from the segregation index for cities with more than 500 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
0.309	64				
2011	0.086	0.134	0.001	0.245	59
2012	0.130	0.266	0.053	0.448	46
2013	0.164	0.194	0.066	0.418	52
2014	0.125	0.105	0.016	0.260	66
2015	0.053	0.216	0.169	0.436	74
2016	0.112	0.121	0.128	0.348	91
2017	0.097	0.155	0.091	0.336	105
2018	0.114	0.170	0.134	0.397	102
2019	0.094	0.147	0.083	0.318	106
2020	0.061	0.128	0.160	0.341	149
all years	0.105	0.160	0.078	0.324	914

Supplementary Table 20: Summary of scaling fits and majority group size adjustment and heterophobia adjustment parameters estimated from the segregation index for cities with more than 500 IAT responses.

year	scaling β_1	majority group size adjustment β_2	$\mathbf{s_1} + \mathbf{s_2}$ β_3	# cities
2010	[-0.094,-0.021]	[-0.303,-0.096]	[-0.135,0.214]	64
2011	[-0.086,-0.010]	[-0.305,-0.073]	[-0.090,0.281]	59
2012	[-0.080,-0.013]	[-0.282,-0.107]	[0.005,0.261]	46
2013	[-0.076,-0.019]	[-0.260,-0.088]	[0.020,0.273]	52
2014	[-0.069,-0.016]	[-0.209,-0.050]	[-0.034,0.226]	66
2015	[-0.047,-0.003]	[-0.210,-0.106]	[0.114,0.290]	74
2016	[-0.055,-0.015]	[-0.174,-0.068]	[0.091,0.270]	91
2017	[-0.059,-0.016]	[-0.210,-0.098]	[0.090,0.302]	105
2018	[-0.061,-0.019]	[-0.238,-0.120]	[0.137,0.354]	102
2019	[-0.063,-0.017]	[-0.229,-0.105]	[0.091,0.331]	106
2020	[-0.044,-0.011]	[-0.183,-0.093]	[0.178,0.362]	149
all years	[-0.045,-0.031]	[-0.179,-0.139]	[0.135,0.201]	914

Supplementary Table 21: Summary of scaling fits and majority group size adjustment and heterophobia variance explained estimated from the gini coefficient for cities with more than 500 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.122	0.180	-0.009	0.309	64
2011	0.086	0.134	0.008	0.251	59
2012	0.130	0.266	0.053	0.448	46
2013	0.164	0.194	0.058	0.413	52
2014	0.125	0.105	0.018	0.261	66
2015	0.053	0.216	0.166	0.433	74
2016	0.112	0.121	0.135	0.354	91
2017	0.097	0.155	0.099	0.343	105
2018	0.114	0.170	0.143	0.405	102
2019	0.094	0.147	0.080	0.316	106
2020	0.061	0.128	0.170	0.350	149
all years	0.105	0.160	0.082	0.327	914

Supplementary Table 22: Summary of scaling fits and majority group size adjustment and heterophobia adjustment parameters estimated from the gini coefficient for cities with more than 500 IAT responses.

year	scaling β_1	majority group size adjustment β_2	$\mathbf{s_1} + \mathbf{s_2}$ β_3	# cities
2010	[-0.094,-0.021]	[-0.305,-0.097]	[-0.118,0.206]	64
2011	[-0.086,-0.010]	[-0.310,-0.077]	[-0.066,0.276]	59
2012	[-0.080,-0.013]	[-0.285,-0.110]	[0.005,0.239]	46
2013	[-0.076,-0.019]	[-0.262,-0.088]	[0.012,0.243]	52
2014	[-0.069,-0.016]	[-0.212,-0.052]	[-0.029,0.207]	66
2015	[-0.047,-0.003]	[-0.213,-0.108]	[0.099,0.256]	74
2016	[-0.055,-0.015]	[-0.177,-0.071]	[0.084,0.240]	91
2017	[-0.059,-0.016]	[-0.212,-0.100]	[0.086,0.270]	105
2018	[-0.061,-0.019]	[-0.241,-0.124]	[0.127,0.315]	102
2019	[-0.063,-0.017]	[-0.232,-0.106]	[0.076,0.283]	106
2020	[-0.044,-0.011]	[-0.187,-0.097]	[0.163,0.321]	149
all years	[-0.045,-0.031]	[-0.182,-0.142]	[0.164,0.241]	914

Supplementary Table 23: Summary of scaling fits and majority group size adjustment and heterophobia variance explained estimated from the η^2 measure for cities with more than 500 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.122	0.180	-0.010	0.309	64
2011	0.086	0.134	-0.001	0.244	59
2012	0.130	0.266	0.043	0.442	46
2013	0.164	0.194	0.062	0.418	52
2014	0.125	0.105	0.015	0.260	66
2015	0.053	0.216	0.140	0.410	74
2016	0.112	0.121	0.091	0.317	91
2017	0.097	0.155	0.097	0.341	105
2018	0.114	0.170	0.116	0.382	102
2019	0.094	0.147	0.056	0.296	106
2020	0.061	0.128	0.115	0.300	149
all years	0.105	0.160	0.056	0.305	914

Supplementary Table 24: Summary of scaling fits and majority group size adjustment and heterophobia adjustment parameters estimated from the η^2 measure for cities with more than 500 IAT responses.

year	scaling β_1	majority group size adjustment β_2	$\mathbf{s_1} + \mathbf{s_2}$ β_3	# cities
2010	[-0.094,-0.021]	[-0.334,-0.091]	[-0.105,0.177]	64
2011	[-0.086,-0.010]	[-0.348,-0.081]	[-0.077,0.224]	59
2012	[-0.080,-0.013]	[-0.329,-0.131]	[-0.004,0.204]	46
2013	[-0.076,-0.019]	[-0.314,-0.118]	[0.014,0.228]	52
2014	[-0.069,-0.016]	[-0.249,-0.059]	[-0.031,0.195]	66
2015	[-0.047,-0.003]	[-0.269,-0.142]	[0.083,0.242]	74
2016	[-0.055,-0.015]	[-0.230,-0.101]	[0.057,0.221]	91
2017	[-0.059,-0.016]	[-0.276,-0.140]	[0.086,0.277]	105
2018	[-0.061,-0.019]	[-0.309,-0.164]	[0.108,0.306]	102
2019	[-0.063,-0.017]	[-0.285,-0.130]	[0.049,0.265]	106
2020	[-0.044,-0.011]	[-0.255,-0.143]	[0.118,0.280]	149
all years	[-0.045,-0.031]	[-0.229,-0.181]	[0.050,0.077]	914

Supplementary Table 25: Comparison of noise ceiling estimates and full sample R^2 for the deviance measure of segregation a threshold of $>$500 responses per city.

year	noise ceiling R^2 range	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	[0.70, 0.92]	0.34	0.48
2011	[0.68, 0.91]	0.26	0.38
2012	[0.53, 0.86]	0.41	0.78
2013	[0.53, 0.86]	0.38	0.72
2014	[0.54, 0.87]	0.25	0.47
2015	[0.37, 0.80]	0.34	0.93
2016	[0.45, 0.84]	0.29	0.63
2017	[0.62, 0.89]	0.30	0.49
2018	[0.60, 0.89]	0.34	0.57
2019	[0.57, 0.88]	0.26	0.47
2020	[0.45, 0.83]	0.25	0.55

Supplementary Table 26: Comparison of noise ceiling estimates and full sample R^2 for the deviance measure of segregation a threshold of $>$250 responses per city.

year	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	0.26	0.47
2011	0.20	0.37
2012	0.23	0.54
2013	0.24	0.70
2014	0.18	0.49
2015	0.22	0.57
2016	0.25	0.69
2017	0.24	0.51
2018	0.33	0.72
2019	0.24	0.55
2020	0.15	0.45

Supplementary Table 27: Comparison of noise ceiling estimates and full sample R^2 for the deviance measure of segregation a threshold of $>$1000 responses per city.

year	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	0.26	0.37
2011	0.31	0.41
2012	0.23	0.41
2013	0.34	0.65
2014	0.39	0.77
2015	0.45	0.79
2016	0.38	0.76
2017	0.35	0.55
2018	0.35	0.50
2019	0.26	0.36
2020	0.30	0.50

Supplementary Table 28: Comparison of noise ceiling estimates and full sample R^2 for the η^2 measure of segregation and a threshold of $>$500 responses per city.

year	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	0.32	0.46
2011	0.26	0.38
2012	0.45	0.86
2013	0.43	0.81
2014	0.27	0.51
2015	0.42	1.13
2016	0.32	0.72
2017	0.35	0.56
2018	0.39	0.65
2019	0.30	0.53
2020	0.30	0.68

Supplementary Table 29: Comparison of noise ceiling estimates and full sample R^2 for the η^2 measure of segregation and a threshold of >250 responses per city.

year	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	0.27	0.49
2011	0.21	0.39
2012	0.24	0.55
2013	0.27	0.79
2014	0.19	0.51
2015	0.25	0.65
2016	0.28	0.75
2017	0.28	0.61
2018	0.37	0.81
2019	0.28	0.64
2020	0.20	0.59

Supplementary Table 30: Comparison of noise ceiling estimates and full sample R^2 for the η^2 measure of segregation and a threshold of >1000 responses per city.

year	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	0.30	0.42
2011	0.32	0.43
2012	0.30	0.54
2013	0.42	0.80
2014	0.43	0.85
2015	0.51	0.90
2016	0.44	0.88
2017	0.42	0.65
2018	0.40	0.58
2019	0.32	0.44
2020	0.37	0.61

Supplementary Table 31: Comparison of noise ceiling estimates and full sample R^2 for the gini coefficient measure of segregation and a threshold of >500 responses per city.

year	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	0.32	0.46
2011	0.26	0.39
2012	0.46	0.87
2013	0.42	0.80
2014	0.27	0.51
2015	0.44	1.19
2016	0.36	0.80
2017	0.35	0.56
2018	0.41	0.68
2019	0.32	0.57
2020	0.35	0.80

Supplementary Table 32: Comparison of noise ceiling estimates and full sample R^2 for the gini coefficient measure of segregation and a threshold of >250 responses per city.

year	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	0.29	0.52
2011	0.22	0.41
2012	0.24	0.55
2013	0.29	0.84
2014	0.19	0.53
2015	0.26	0.69
2016	0.32	0.87
2017	0.31	0.67
2018	0.39	0.86
2019	0.33	0.75
2020	0.26	0.77

Supplementary Table 33: Comparison of noise ceiling estimates and full sample R^2 for the gini coefficient measure of segregation and a threshold of >1000 responses per city.

year	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	0.31	0.44
2011	0.33	0.45
2012	0.34	0.62
2013	0.43	0.82
2014	0.43	0.84
2015	0.51	0.89
2016	0.44	0.88
2017	0.43	0.66
2018	0.41	0.58
2019	0.33	0.46
2020	0.40	0.67

Supplementary Table 34: Comparison of noise ceiling estimates and full sample R^2 for the segregation index measure of segregation and a threshold of >500 responses per city.

year	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	0.32	0.45
2011	0.26	0.38
2012	0.46	0.87
2013	0.43	0.81
2014	0.27	0.51
2015	0.44	1.20
2016	0.36	0.78
2017	0.34	0.55
2018	0.40	0.67
2019	0.32	0.57
2020	0.35	0.78

Supplementary Table 35: Comparison of noise ceiling estimates and full sample R^2 for the segregation index measure of segregation and a threshold of >250 responses per city.

year	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	0.28	0.50
2011	0.22	0.41
2012	0.24	0.55
2013	0.28	0.83
2014	0.19	0.52
2015	0.26	0.68
2016	0.31	0.85
2017	0.29	0.64
2018	0.39	0.85
2019	0.32	0.74
2020	0.25	0.74

Supplementary Table 36: Comparison of noise ceiling estimates and full sample R^2 for the segregation index measure of segregation and a threshold of >1000 responses per city.

year	Full Sample R^2	Lower Bound Noise Corrected R^2
2010	0.31	0.44
2011	0.33	0.44
2012	0.34	0.62
2013	0.44	0.83
2014	0.43	0.84
2015	0.52	0.92
2016	0.45	0.90
2017	0.42	0.66
2018	0.41	0.58
2019	0.34	0.47
2020	0.39	0.65

Supplementary Table 37: Summary of scaling fits and majority group size adjustment and heterophobia adjustment variance explained for cities with more than 250 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.125	0.128	-0.007	0.255	126
2011	0.056	0.120	-0.005	0.188	119
2012	0.165	0.047	-0.008	0.225	88
2013	0.134	0.091	-0.008	0.232	98
2014	0.097	0.066	-0.008	0.172	116
2015	0.044	0.158	-0.004	0.211	129
2016	0.078	0.145	0.021	0.246	148
2017	0.079	0.132	0.015	0.232	163
2018	0.143	0.131	0.061	0.323	157
2019	0.054	0.168	0.008	0.236	174
2020	0.046	0.059	0.040	0.150	228
all years	0.086	0.117	0.003	0.205	1546

Supplementary Table 38: Summary of scaling fits and majority group size adjustment and heterophobia adjustment variance explained for cities with more than 1000 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.053	0.143	-0.022	0.239	31
2011	0.112	0.145	-0.030	0.284	31
2012	0.047	0.082	-0.019	0.195	26
2013	0.052	0.195	0.016	0.319	28
2014	0.039	0.262	0.036	0.371	34
2015	0.014	0.340	0.055	0.436	43
2016	0.115	0.227	0.022	0.369	57
2017	0.071	0.207	0.045	0.339	61
2018	0.090	0.210	0.022	0.337	59
2019	0.054	0.169	-0.002	0.248	61
2020	0.041	0.147	0.094	0.292	89
all years	0.076	0.201	0.026	0.298	520

Supplementary Table 39: Summary of scaling fits and majority group size adjustment and heterophobia adjustment variance explained from the η^2 measure for cities with more than 250 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.125	0.128	0.010	0.265	126
2011	0.056	0.120	0.012	0.199	119
2012	0.165	0.047	0.002	0.228	88
2013	0.134	0.091	0.035	0.263	98
2014	0.097	0.066	0.004	0.180	116
2015	0.044	0.158	0.033	0.243	129
2016	0.078	0.145	0.050	0.270	148
2017	0.079	0.132	0.071	0.278	163
2018	0.143	0.131	0.118	0.368	157
2019	0.054	0.168	0.052	0.274	174
2020	0.046	0.059	0.094	0.199	228
all years	0.086	0.117	0.036	0.231	1546

Supplementary Table 40: Summary of scaling fits and majority group size adjustment and heterophobia adjustment variance explained from the η^2 measure for cities with more than 1000 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.053	0.143	0.024	0.276	31
2011	0.112	0.145	-0.008	0.299	31
2012	0.047	0.082	0.069	0.269	26
2013	0.052	0.195	0.115	0.401	28
2014	0.039	0.262	0.085	0.413	34
2015	0.014	0.340	0.126	0.501	43
2016	0.115	0.227	0.104	0.434	57
2017	0.071	0.207	0.125	0.407	61
2018	0.090	0.210	0.091	0.393	59
2019	0.054	0.169	0.065	0.306	61
2020	0.041	0.147	0.167	0.358	89
all years	0.076	0.201	0.088	0.351	520

Supplementary Table 41: Summary of scaling fits and majority group size adjustment and heterophobia adjustment variance explained from the segregation index for cities with more than 250 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.125	0.128	0.023	0.274	126
2011	0.056	0.120	0.024	0.210	119
2012	0.165	0.047	0.006	0.231	88
2013	0.134	0.091	0.053	0.277	98
2014	0.097	0.066	0.009	0.183	116
2015	0.044	0.158	0.045	0.253	129
2016	0.078	0.145	0.092	0.307	148
2017	0.079	0.132	0.085	0.290	163
2018	0.143	0.131	0.137	0.382	157
2019	0.054	0.168	0.103	0.320	174
2020	0.046	0.059	0.148	0.250	228
all years	0.086	0.117	0.064	0.255	1546

Supplementary Table 42: Summary of scaling fits and majority group size adjustment and heterophobia adjustment variance explained from the segregation index for cities with more than 1000 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.053	0.143	0.040	0.288	31
2011	0.112	0.145	0.004	0.307	31
2012	0.047	0.082	0.121	0.312	26
2013	0.052	0.195	0.136	0.418	28
2014	0.039	0.262	0.079	0.408	34
2015	0.014	0.340	0.134	0.508	43
2016	0.115	0.227	0.113	0.441	57
2017	0.071	0.207	0.134	0.414	61
2018	0.090	0.210	0.092	0.395	59
2019	0.054	0.169	0.085	0.324	61
2020	0.041	0.147	0.195	0.383	89
all years	0.076	0.201	0.111	0.372	520

Supplementary Table 43: Summary of scaling fits and majority group size adjustment and heterophobia adjustment variance explained from the gini coefficeint for cities with more than 250 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.125	0.128	0.031	0.281	126
2011	0.056	0.120	0.027	0.212	119
2012	0.165	0.047	0.007	0.231	88
2013	0.134	0.091	0.059	0.282	98
2014	0.097	0.066	0.009	0.184	116
2015	0.044	0.158	0.050	0.258	129
2016	0.078	0.145	0.100	0.313	148
2017	0.079	0.132	0.104	0.306	163
2018	0.143	0.131	0.147	0.391	157
2019	0.054	0.168	0.106	0.323	174
2020	0.046	0.059	0.158	0.260	228
all years	0.086	0.117	0.067	0.258	1546

Supplementary Table 44: Summary of scaling fits and majority group size adjustment and heterophobia adjustment variance explained from the gini coefficeint for cities with more than 1000 IAT responses.

year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.053	0.143	0.044	0.291	31
2011	0.112	0.145	0.010	0.310	31
2012	0.047	0.082	0.120	0.311	26
2013	0.052	0.195	0.128	0.410	28
2014	0.039	0.262	0.079	0.408	34
2015	0.014	0.340	0.118	0.494	43
2016	0.115	0.227	0.105	0.434	57
2017	0.071	0.207	0.136	0.415	61
2018	0.090	0.210	0.093	0.395	59
2019	0.054	0.169	0.078	0.317	61
2020	0.041	0.147	0.205	0.393	89
all years	0.076	0.201	0.111	0.371	520