THE UNIVERSITY OF CHICAGO

URBAN PSYCHOLOGY: A FRAMEWORK FOR ECOLOGICALLY RELEVANT INVESTIGATIONS OF HUMAN BEHAVIOR

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY
ANDREW JACOB STIER

CHICAGO, ILLINOIS
AUGUST 2023

To my advisors, who inspired my fascination with cities and fostered and encouraged intellectual creativity.

It may be romantic to search for the salve of society's ills in slow-moving rustic surroundings, or among innocent, unspoiled provincials, if such exist, but it is a waste of time. Does anyone suppose that, in real life, answers to any of the great questions that worry us today are going to come out of homogeneous settlements?

-Jane Jacobs

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Are people more or less depressed in larger cities? Are attention spans shorter in busy urban areas? Urban Psychology is the study of how the built environment of cities influences human behavior and causes psychological adaptations at the individual level. In this dissertation, I present research extending Urban Scaling Theory models to better understand how cities shape human psychology. I discuss (1) an application of these models to reveal how cities systematically influence the risk of psychological depression, (2) an extension of these models to understand how heterogeneous interactions in cities influence implicit racial bias levels and economic outputs, and (3) evidence that selective attention capabilities are increased in larger cities. These studies are first forays into Urban Psychology research and pave the way for future projects that aim to determine the interplay between individual and neighborhood characteristics and average behaviors for entire urban areas. My hope is that the field of urban psychology brings more psychological focus to understanding how cities' complex physical and social environments constrain behavior through regularities of human mobility but also provide a greater diversity of social and economic experiential options. Doing so will hopefully lead to new discoveries in terms of understanding human behavior and the inner workings of large-scale complex systems.

CHAPTER 1

INTRODUCTION

We, human beings, are continuously modifying the environment to suit our needs. We harvest plants for food, build roads, and turn metals into useful tools which enhance our ability to modify the environment, in turn. On longer timescales, we select for tastier carrots, breed more docile livestock, and change entire climates, geographies, and ecosystems to suit our needs. Arguably, the most impressive and impactful of these environmental modifications is the city.

Cities provide opportunities for more efficient uses of resources and greater possibilities for social interactions. However, at least historically, cities have also facilitated more crime, disease, pollution, and exploitation. As we shall see, many of the various benefits and drawbacks that cities provide result precisely from the interactions that cities facilitate. From this perspective of extensive environmental modification, it is not so strange to claim that as we change our surroundings our psychology changes: we are engaged in a constant process of co-determination with the environment Clark [2015], Berman et al. [2019].

The research presented here is in the spirit of this claim. It attempts to explicitly recognize that cities do something more profound than simply providing different modes for us to interact with the built environment and other people. Cities can, and often do, engender psychological adaptations in their inhabitants. The agency given to cities here is purposeful: the observed patterns of behavior in cities are co-determined by features of the built environment and the psychological adaptations of their inhabitants to the social and physical niches that are available.

These are strong claims. The research presented here does not prove them definitively. However, it does make a strong case when situated within the context of a large body of ecological (e.g., see the overviews in Clark [2015] and Henrich [2016] for examples from gene-culture co-evolution) and urban science literature Bettencourt [2021a] on niche

construction (i.e., this process of co-determination) and adaptation. These fields have laid the groundwork for Urban Psychology by painstakingly investigating similar phenomena in individual organisms and in small-scale, non-modernized societies. Urban Psychology, in contrast, extends these ideas to the tens of millions of people that live in the largest scale human settlements, i.e., cities.

Each of the chapters that follow serves to provide additional building blocks for a general theory of how cities arise and shape their inhabitants. Specifically, in Chapter 2, I demonstrate how the increased social network density of larger cities can provide an environment that buffers against psychological depression. In Chapters 3 and 4 I demonstrate the effects of demographics and residential segregation on economic outputs (gross domestic product and incomes) and implicit racial biases. Finally, in Chapter 5 I provide evidence that some aspects of psychological attention may be improved in larger cities and that this may be related to evidence that social networks are less tightly knit in large cities.

It is certainly possible to study these psychological and behavioral phenomena in other settings (e.g., in a psychology laboratory). Indeed, I draw heavily on experiments and evidence from more traditional controlled environments throughout this research. However, those settings are typically controlled to the point that it is all but impossible to study complex interactions between the environment (social and physical) and human psychology. Thus, the research that follows is also an attempt to, in a principled and theoretically grounded manner, bring psychological research into real-world settings that are particularly relevant for modern-day humans. At the same time, this research is a continuation and extension of urban thought that has long existed outside of psychology (e.g., see Jacobs [2016] and Bettencourt [2021a]).

Before we continue, I find it helpful to have a definition for Urban Psychology. The one I propose is that "Urban Psychology is the study of human behavior through the lens of feedback loops between the structured environments of cities and human cognition." I

like this definition for two reasons. First, it makes clear what we should focus on when we do Urban Psychology research. Specifically: human behavior, structured environments of cities, and feedback loops with a particular emphasis on cognition. Second, this definition is suggestive of the fact that cities are a specific example of complex systems (a system with many parts that shows non-trivial collective behavior Newman [2011]) which have structured environments that constrain agents and have agents that can dynamically adapt to the environment and directly modify it. Finally, a number of goals for Urban Psychology research flow naturally (in increasing levels of generality) from this definition: first, to help us better understand cities, second, to better understand human behavior in general, and third, to prove helpful towards research into similar classes of complex systems, be they in non-human animals or inorganic matter.

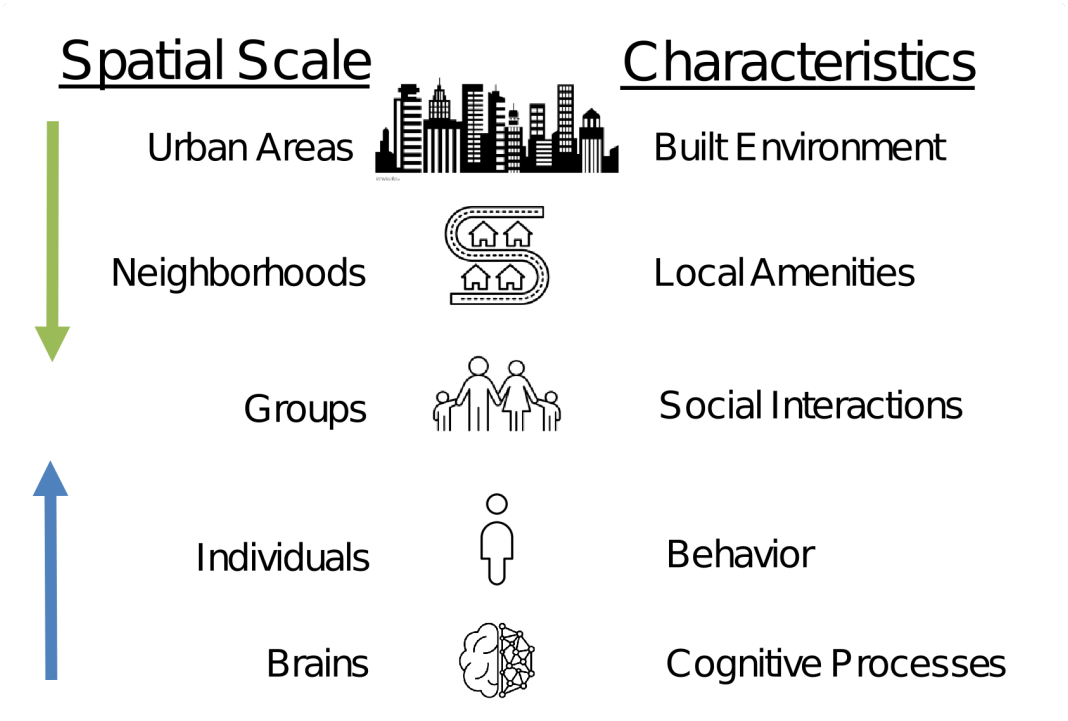


Figure 1.1: My approach to urban psychology is to unite approaches from urban science which primarily have focused on large-scale phenomena at the level of entire urban areas and neighborhoods with approaches from psychology which have primarily focused on individual behavior and how cognitive processes are instantiated in the brain.

The specific approach that is taken in the research that follows is to rely heavily on existing approaches for studying cities and then integrate psychological research questions and methods. From the point of view of the various spatial scales that exist within cities, my general approach has been to integrate mathematical models and tools which capture behavior at the organizational level of entire cities with existing and new research that investigates behavior at the level of individual inhabitants (Figure 1.1). At the top of this hierarchy are the tools of modern urban science Bettencourt [2021a], which tend to focus more on the types of things that people produce in cities, For example, innovations, economic outputs, and infrastructure networks. At the bottom of this hierarchy are the tools of traditional psychological research that have focused much more on individual behavior, cognition, and how cognition is instantiated in the brain.

I will close by mentioning that the foundation of all of the research that follows is seminal work led by Luís Bettencourt Bettencourt [2013, 2021b] which recognized that cities are amenable to modeling with tools from the field of complex systems. Undoubtedly, this research, which is known as the Urban Scaling Theory framework, drew on a long literature of research in urban planning, economics, and sociology that had described many aspects of city life and organization in detail. However, under the complex systems framework, Urban Scaling Theory demonstrated that much of the intricate detail of city life, about neighborhoods, groups, and individuals, can be ignored while still learning quite a bit about how and why cities function. In fact, it turns out that the top-down models of average behavior for entire urban areas that form the basis for the research that follows, are particularly informative with regard to mechanisms of human behavior in cities. For example, some of the initial results of Urban Scaling theory were the demonstrations that cities exhibit economies of scale in infrastructure networks and a densification of social networks as cities grow more populous.

In summary, the research presented here presents a novel framework for understanding

human behavior in ecological settings. While there is still much to discover, these efforts have already provided insight into psychological depression (Chapter 2), racial segregation (Chapter 3), implicit racial bias (Chapter 4), and selective attention (Chapter 5).

CHAPTER 2

EVIDENCE AND THEORY FOR LOWER RATES OF DEPRESSION IN LARGER U.S. URBAN AREAS[1]

It is commonly assumed that cities are detrimental to mental health. However, the evidence remains inconsistent and, at most, makes the case for differences between rural and urban environments as a whole. Here, we propose a model of depression driven by an individual's accumulated experience mediated by social networks. The connection between observed systematic variations in socioeconomic networks and built environments with city size provides a link between urbanization and mental health. Surprisingly, this model predicts lower depression rates in larger cities. We confirm this prediction for US cities using four independent datasets. These results are consistent with other behaviors associated with denser socioeconomic networks and suggest that larger cities provide a buffer against depression. This approach introduces a systematic framework for conceptualizing and modeling mental health in complex physical and social networks, producing testable predictions for environmental and social determinants of mental health also applicable to other psychopathologies.

Introduction

Living in cities changes the way we behave and think Milgram [1970b], Simmel [2012], Bettencourt [2021a]. Over a century ago, the social changes associated with massive urbanization

1. This Chapter was published as: Stier, Andrew J., Kathryn E. Schertz, Nak Won Rim, Carlos Cardenas-Iniguez, Benjamin B. Lahey, Luís MA Bettencourt, and Marc G. Berman. "Evidence and theory for lower rates of depression in larger US urban areas." Proceedings of the National Academy of Sciences 118, no. 31 (2021): e2022472118.

in Europe and in the United States, focused social scientists on the nexus between cities and mental life Simmel [2012]. Along with the urban public health crises of the time, a central question became whether cities are good or bad for mental health.

Subsequently, social psychologists Milgram [1970b] started to document and measure the systematic behavioral adaptations among people living in cities. These adaptations included strategies to curb unwanted social interactions – such that people in larger cities act in colder and more callous ways Milgram [1970b], a more intense use of time (e.g. faster walking Bettencourt et al. [2007a]), and a greater tolerance for diversity Wilson [1985]. These studies attributed the influences of urban environments on mental health to the intensity of social life in larger cities, mediated by densely built spaces and associated dynamic and diverse socioeconomic interaction networks. They did not, however, ultimately clarify whether urban environments promote better or worse mental health. Consequently, concerns persisted that cities are mentally taxing Lewis and Booth [1994], Gruebner et al. [2017], Krabbendam and Van Os [2005], Sundquist et al. [2004] and can induce "stimulus overload", including stress, mental fatigue Berman et al. [2008], and low levels of subjective well-being Morrison and Weckroth [2018].

More recent studies have focused less on urban environments as a whole and more on contextual and environmental factors associated with depression. For example, a study of the entire population in Sweden Sundquist et al. [2004] uncovered a positive association between neighborhood population density and depression-related hospitalizations. In addition, individual factors of gender, age, socioeconomic status, and race, which vary at neighborhood levels within cities, have been found to be statistically associated with depression Blanco et al. [2010], Mirowsky and Ross [2001], Gilman et al. [2003]. Other studies using various measures of mental health and broader definitions of urban environments have found evidence for an association between poorer mental health in cities versus rural areas Gruebner et al. [2017], Krabbendam and Van Os [2005]. However, this evidence, and that linking sub-

jective well-being and cities Glaeser et al. [2016], Itaba [2016], Mitchell et al. [2013], Chen et al. [2015], has remained mixed and often explicitly inconsistent Lee [2014], Kearns et al. [2012], due to differences in: 1) reporting (e.g., surveys vs. medical records), 2) types of measurement (e.g., surveys vs. interviews), 3) definitions of what constitutes urban, and 4) the mental disorders studied (e.g. schizophrenia vs. depression).

For these reasons, it is desirable to create a systematic framework that organizes this diverse body of research and interrogates how varying levels of urbanization influence mental health across different sets of indicators. Here, we begin to build this framework for depression in US cities. We show that, surprisingly, the per capita prevalence of depression decreases systematically with city size.

Like earlier classical approaches, our strategy frames the effects of city size on mental health through the lens of the individual experience of urban physical and socioeconomic environments. Crucial to our purposes, many characteristics of cities have been recently found to vary predictably with city population size. These systematic variations in urban indicators are explained by denser built environments and their associated increases in the intensity of human interactions and resulting adaptive behaviors Bettencourt [2013].

More specifically, people in larger cities have, on average, more socioeconomic connections mediating a greater variety of functions. This effect is understood theoretically by the statistical likelihood to interact with more people over space per unit time, leading to both potential mental "overload" but also to greater stimulation and choice along more dimensions of life. This expansion of socioeconomic networks is supported structurally by economies of scale (e.g. road length) in urban built environments and by occupational specialization and associated increases in economic productivity and exchange Bettencourt [2021a].

This effect leads to a number of quantitative predictions about the nature of urban spaces and socioeconomic variables, the most central of which is the variation of the average number of socioeconomic interactions, k (network degree) with city size, N, as $k(N) = k_0 N^{\delta} e^{\xi}$.

Here, k_0 is a prefactor independent of city size, and ξ a residual measuring the distance from the population average. The exponent $0 < \delta \simeq 1/6 < 1$ measures the percent increase in the number of connections with each percent increase in city population, which is an elasticity in the language of economics. Because the ξ reflects city-size independent statistical fluctuations, these errors average out across cities and k obeys a scaling relationship on average over cities, such that $k(N) \sim N^{\delta}$. This expectation is directly observed in cell phone networks Schläpfer et al. [2014a] and indirectly via the faster spread of infectious diseases such as COVID-19 Stier et al. [2020], and by higher per capita economic productivity and rates of innovation Bettencourt et al. [2007a], Bettencourt [2013].

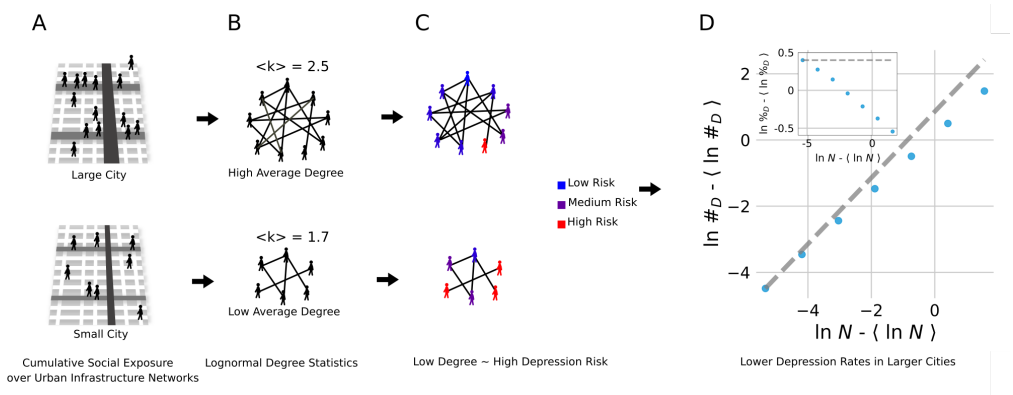


Figure 2.1: Sub-linear scaling of depression in a social network model. (a) Individuals moving over a city's hierarchical infrastructure network experience cumulative exposure to semi-random social interactions. (b) This cumulative exposure results in social networks with log-skew-normal degree (k) statistics with a mean which increases with city size indicating more per-capita social interactions in larger cities, on average. (c) Individual risk for depression is inversely proportional to social connectivity (degree) and is superimposed on the social networks generated within cities (d) The combination of how cities shape social networks and how social networks shape individual depression risk results in a prediction of sub-linear scaling of depression cases with increased city size, i.e., lower depression rates in larger cities (Inset). The logarithm of population and depression incidence are mean centered for ease of comparison to the empirical results.

This result is important to mental health because depression is associated, at the individual level, with fewer social contacts Okamoto et al. [2011], Rosenquist et al. [2011]. To translate the general scaling of social interactions with city size into a model for the incidence

of depression in urban areas, we will now need to pay particular attention not only to the average number of social connections in a city of size N, $k(N)$, but also to its variance across individuals in that city and how they influence depression.

Results

We developed a statistical mathematical model that brings together socioeconomic network structure with individual risk of depression (Fig. 2.1). This model takes the form of a generative social network, which combines: i) a degree distribution with mean scaling as $k(N) = k_0 N^{\delta}$ (Fig. 2.1B) with ii) the risk (probability) for an individual to manifest depression, $p_d(k)$, taken to be inversely proportional to their social connectivity, $p_d(k) \sim 1/k$ (Fig. 2.1C). We will return to the finer issue of quality and type of connections below. For now, note that a larger number of connections in larger cities entails a qualitatively different experience because it is driven by the need to obtain support, goods, and services in environments with deep divisions of knowledge and labor.

To complete the model, we need to specify the probability distribution of degree, $f(k)$ in each city. We adopt a log-skew-normal distribution with parameters similar to those measured in Schläpfer et al. [2014a], see Fig. 2.1B. This choice introduces another assumption into our model because lognormal distributions arise from multiplicative random processes, which compound risk over time to generate outcomes. In this sense, the adoption of this distribution assumes that depression is the result of a cumulative exposure process over time Vinkers et al. [2014], see Fig. 2.1A, mediated by an individual's social network. Fig. 2.1D shows results from this model obtained by sampling each city's degree distribution N times, corresponding to a city's population. Each simulated city resident is then diagnosed with a binary outcome, manifesting depression or not proportionally to their individual risk, $p_d(k)$.

We used this model to generate urban socioeconomic networks and computed their as-

sociated number of depression cases, Y, for a range of city sizes from $N = 10^4 - 10^7$ that span population range of US metropolitan areas, see Fig. 2.1D. We observed a simple scaling relation for the total number of depressive cases

$$Y(N_i, t) = Y_0(t)N_i(t)^{\beta}e^{\xi_i(t)}. \tag{2.1}$$

with a sub-linear exponent $\beta = 1 - \delta < 1$. For $\beta = 1$ ($\delta = 0$), cases of depression increase proportionally to population so that there would be no city size effect. In contrast, for $\beta < 1$ (sub-linear), a smaller proportion of the population manifests depression in larger cities.

We express the quantitative consequences of the model based on 100 iterations for each city to predict that the number of depression cases follows a power law function of city size with a scaling exponent $\beta = 0.859$ (95% CI = [0.854, 0.863]), Fig. 2.1D. Thus, under the model's assumptions, we expect larger cities to show substantially lower per capita rates of depression.

To test these quantitative expectations, we asked whether empirical measurements of depression exhibit a systematic scaling relationship with city population size. We analyzed four independent data sets, which allow for consistent assessments of cases of depression across different urban areas in the US.

First, we employed estimates of the prevalence of depression in US cities produced as a part of two annual population surveys: the National Survey on Drug Use and Health (NSDUH) Abuse and Administration [2011] from the Substance Abuse and Mental Health Services Administration (SAMHSA) and the Behavioral Risk Factor Surveillance System (BRFSS) from the Centers of Disease Control (CDC), see Methods, Fig. S1, S3, and Tables S2, S3, S4. The NSDUH asks respondents whether they have experienced a major depressive episode in the past year, as defined by the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) Abuse and Administration [2011]. The BRFSS asks respondents if they have ever been told that they have a depressive disorder. Both surveys involved a social

interaction between a surveyor and the respondent, which takes place over the phone for the BRFSS and in person for the NSDUH. The differences between the two surveys provide a consistency test on measured cases of depression and partially rule out the possibility that their variation with city size is idiosyncratic to particular experimental or survey methodologies.

Second, to generalize across different indicators and to avoid biases in reporting due to social stigma Mak et al. [2007], we added two additional estimates of depression prevalence based on passive observation, which does not rely on an overt survey instrument. Specifically, we explored two large geo-located Twitter datasets of individuals and their messages for depressive symptoms in different cities. Twitter requires users to opt-in to geo-location and as a result only a small fraction of tweets are geolocated Cuomo et al. [2021]. Importantly, this bias further distinguishes the two Twitter datasets from the survey-based data and strengthens any claims of generalization across the ways in which the data were collected and the populations of people studied.

These two Twitter datasets included an existing dataset collected over one week in 2010 Eisenstein et al. [2010] and a new historical dataset covering one month in 2019. Similar datasets have been used to demonstrate that happiness decreases with per-capita tweets Mitchell et al. [2013], that counts of users scale super-linearly with city size Arthur and Williams [2019], and to assess regional variability in subjective well-being Iacus et al. [2019], but to our knowledge they have not been used to directly estimate associations between mental health disorders and city size.

To measure the prevalence of depression from this corpus, we employed a machine learning technique to identify depressive symptoms from users' messages, emulating the Patient Health Questionnaire-9 (PHQ-9) commonly used by clinicians. The PHQ-9 consists of 9 questions based on the nine criteria for diagnosing depression in the DSM-IV. In order to emulate the PHQ-9 questions, we used a previously determined lexicon of seed terms orga-

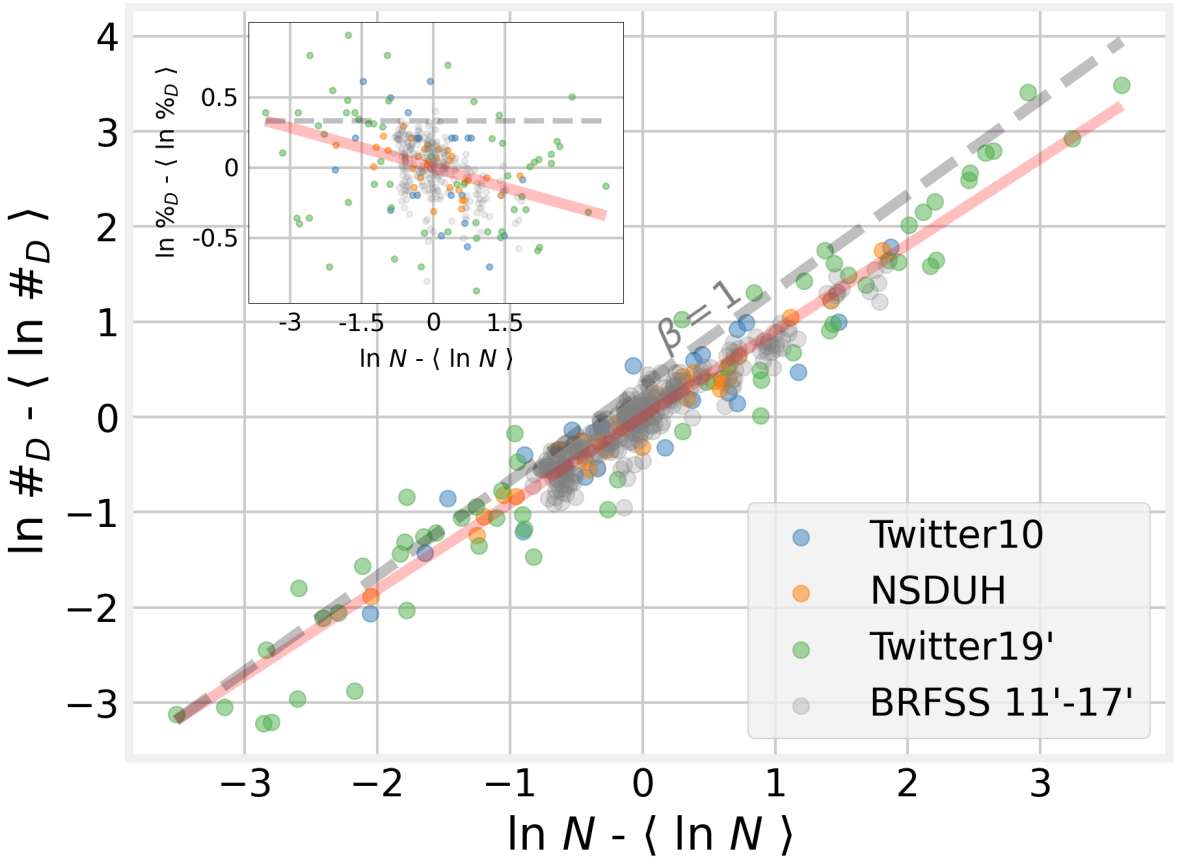


Figure 2.2: Depression cases scale sub-linearly with city size. City level measures of depression prevalence were obtained from two survey based data sets (NSDUH and BRFSS) and two passive observation data sets (Twitter10' and Twitter19'). To collapse across datasets the natural log of Population, N, and estimated total depression cases, $\#_D$, were mean-centered within each dataset. An ordinary least squares linear regression of the pooled data resulted in an estimate of $\beta = 0.904$, 95% CI = [0.853, 0.956], and an R^2 of 0.75. Inset: depression rates decrease with city size. $\beta = -0.096$, 95% CI = [-0.12, -0.07], $R^2 = 0.13$.

nized into nine topics to guide a Latent Dirichlet Allocation Yazdavar et al. [2017] method to determine the degree to which each user's messages represent these topics, see Methods, Fig. S4, Table S5. This technique has been found to have an accuracy (proportion of tweets correctly identified) of 68% and precision (1 - the false discovery rate of tweets with depressive symptoms) of 72% compared to expert assignment of tweets to PHQ-9 questions Yazdavar et al. [2017].

We estimated the scaling exponent β from each of these datasets via ordinary least squares (OLS) linear regression between the logarithm of total depression cases and the logarithm of population size (See Methods, Figs. S4 and S5, Table S6). When pooling across datasets and years we estimated a scaling exponent of $\beta = 0.904$ (95% Confidence Interval (CI) = [0.853, 0.956]) (Fig. 2.2), consistent with our simulation model's prediction of $\beta = 0.859$. Moreover, estimates of β are similar when calculated separately for each dataset (Table 2.1, Fig. S6).

While the Twitter19' dataset suggests that this statistical relationship is consistent across cities with populations ranging from about 15 thousand to 20 million (Table S1), the BRFSS dataset only supports sublinear scaling of depression rates in cities larger than $\sim$0.5 million people (see Methods, Fig. S1). This discrepancy may be due to the fact that the BRFSS city level data is only reported for cities with at least 500 respondents in order to ensure anonymity. We provide evidence that this cutoff artificially alters the joint distribution of depression prevalence estimates and city size in the BRFSS data but, importantly, we find no evidence of similar non-linear shifts in the joint distribution in the Twitter19' dataset (see Methods, Figures S1, S2).

As an additional sensitivity analysis we performed a logistic regression to assess how conditioning on race, income, education, and rate of population change (i.e. migration) impacted the observed decrease in depression rates for larger cities. We did this with individual level survey responses for each year of the BRFSS data. Similar to the scaling analysis above,

the average odds ratio across all years for a one unit increase in the natural logarithm of city population was .89 (maximal 95% CI [.87,.93], see Methods, Table S7). Population change was also not significantly related to depression rate in the NSDUH and Twitter datasets (Tables S8). Thus we find general empirical support for the expectation that larger cities are associated with a decreased risk for depression even when conditioning on race, education, income, and migration. We found no consistent evidence that the rate of population change (i.e. migration rate) was associated with depression rates across all datasets, despite previous research associating growing cities with increased subjective well-being Glaeser et al. [2016].

This statistical relationship between depression and city size is consistent in larger cities across all four datasets and across a decade, despite the different ways in which depressive symptoms are measured and the different ways that the data were collected. Importantly, these results demonstrate that depression rates are substantially lower in larger US cities, contrary to previous expectations, but precisely in line with our theoretical model and simulations.

Table 2.1: Estimates of the scaling exponent β for each dataset. In all cases we observe sub-linear ($\beta < 1$) scaling of total depression cases with city size. n indicates the number of cities included in each dataset.

Dataset	β	95% CI	R^2	n
Twitter10'	0.822	[0.671,0.973]	0.853	24
NSDUH	0.887	[0.826,0.949]	0.968	31
Twitter19'	0.942	[0.886,0.998]	0.951	60
BRFSS2011	0.881	[0.778,0.983]	0.881	43
BRFSS2012	0.854	[0.741,0.966]	0.865	39
BRFSS2013	0.860	[0.750,0.970]	0.865	41
BRFSS2014	0.829	[0.737,0.922]	0.902	38
BRFSS2015	0.818	[0.733,0.902]	0.903	43
BRFSS2016	0.827	[0.746,0.907]	0.913	43
BRFSS2017	0.832	[0.769,0.896]	0.949	40

Discussion

Although the association between urbanization and mental health is foundational in the social sciences and in public health, it has remained challenging to characterize and assess quantitatively. This is particularly concerning as almost every nation worldwide continues to urbanize, with over 70% of the world's population expected to live in cities by 2050 United Nations and Social Affairs [2019], and depressive disorders already a leading global cause of disability James et al. [2018] and economic losses Greenberg et al. [2003].

Based on size alone, large cities bear the brunt of the social and economic burden of depressive disorders. Our findings suggest that on a relative basis, however, smaller cities are actually worse off. Consequently, the discrepancy between BRFSS data and Twitter19' data in small cities is particularly concerning. While our analysis suggests that this discrepancy stems from the way in which BRFSS city level data are reported, it is important that future work develops accurate observation instruments for both smaller cities and finer geographic units within cities (i.e. neighborhoods). This will be particularly important as public health officials start to incorporate geographic patterns of mental health disorders into their allocation plans for mental health care resources.

The convergence of recent findings from urban science with evidence and theory from mental health studies offers a window for creating more systematic approaches to understanding mental health in cities. In this respect, the sub-linear scaling of total depression cases with population size in larger U.S. cities is a completely unexpected result characterizing the socio-geographic distribution of depression. While the results presented here speak only to larger urban areas in the U.S., they suggest that larger city environments and urbanization can, on average, naturally provide greater social stimulation and connections that may buffer against depression. Though urban scaling theory has been shown to generalize across cultures Bettencourt [2013] and human history Ortman et al. [2015, 2016, 2014a], Bettencourt [2021a], it is critical for future work to examine whether the presented extension

of urban scaling theory to depression generalizes to smaller cities and to other countries and cultures.

While our theoretical model only considers the quantity of social connections, embedded in urban scaling theory is the general implication that the net (economic and social) benefit of these interactions is positive Bettencourt [2013]. Future work on the link between social connectivity and mental health should consider and explicitly model urban gradients in the quality of such connections. Alongside quantity, the quality of social connections is a strong predictor of depression Seo et al. [2015], Werner-Seidler et al. [2017] and subjective well-being (SWB) Seo et al. [2015]. Though the evidence relating subjective well-being to cities is mixed – some studies report no relationship between city size and SWB Glaeser et al. [2016], some suggest higher SWB in larger cities Itaba [2016], some suggest lower SWB in larger cities Mitchell et al. [2013], and some suggest an inverted U, i.e., higher SWB for mid-sized cities Dang et al. [2020], Chen et al. [2015] – the quality of social connections might hold the key to understanding discrepancies between city level trends in SWB and depression rates.

In particular, while positive and negative affect are similarly weighted in subjective well-being measures Burns et al. [2011], depression is frequently characterized by more substantial and nuanced changes in negative affect Demiralp et al. [2012], Panaite et al. [2020]. Thus, the results presented here suggest that the greater number of social connections in larger cities on the whole may provide a social buffer against negative affect and depression in the most vulnerable people (i.e. those with the smallest social networks).

Conversely, increased positive affect is related to higher quality social connections independently of negative affect and depressed mood Steptoe et al. [2009]. Thus, the alleged more callous and superficial social interactions in larger cities Milgram [1970b], Bettencourt et al. [2007a] may explain decreases in positive affect and subjective well-being; but simultaneously may still buffer individuals from depression by decreasing negative affect (i.e.

these more numerous social interactions may impact negative and positive affect differently). Since, individuals with the lowest subjective well-being are at a significant risk for clinical depression Diener and Michalos [2009], Koivumaa-Honkanen et al. [2004], Seo et al. [2015], Oswald and Wu [2011], it is important for future work to examine how the rate of low SWB scores varies between cities of different sizes.

We must also recognize that the numerous factors which influence depression vary enormously within cities. These variations may influence individuals directly and also indirectly through the local environments in which they live and work.

For example, homophilic gradients of mobility have been observed in neighborhoods with similar levels of socio-economic status (SES) Saxon [2020], Graif et al. [2017b], Lathia et al. [2012b], so that city inhabitants from poorer (richer) communities tend to preferentially travel to similarly poor (rich) areas. In addition, recent research suggests that neighborhoods with higher overall socio-economic status tend to be better integrated into their surroundings, affording residents better access to the rest of the city Saxon [2020]. Thus, it is crucial that future work examines the relationship between depression rates, mobility, and social connectivity in smaller populations, such as at the neighborhood level.

In addition, looking within cities at these local and more fine-grained levels is expected to reveal variations in the incidence of depression via other social groupings Blanco et al. [2010]. For example, several studies have associated high population density in social housing in Europe and the US with higher incidence of depression in aging adults Lee [2014], possibly mediated by a higher density of negative connections with neighbors, which can instill feelings of isolation, fear and despair. In order to search for finer causal evidence, future work may employ a number of experimental designs such as sibling comparisons or stratification by confounding factors D'Onofrio et al. [2020].

Examining scaling relationships of mental health outcomes with city size is a systematic way of investigating general urban effects on mental life which places focus on collective

influence on mental health disorders. The perspective of cities as interconnected networks which shape their inhabitants lives may also help to uncover environmental factors that influence other mental health disorders and overall well-being. This includes highly co-morbid psychopathologies such as anxiety disorders, and less co-morbid ones such as schizophrenia, for which increased socialization may lead to different outcomes. The fact that important insights about the mechanisms of mental health disorders might be gleaned from such a general population level analysis, which ignores the intricate and often personal details of mental health, is surprising and powerful.

Materials and Methods

Data Sources and Processing

County populations in Fig. 2.2 are provided by the United States Census Bureau and available online at `https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html`. We used delineation files provided by the US Office of Budget and Management to aggregate county level data up to Metropolitan Statistical Areas (MSAs). Each MSA is represents a US Census definition of a functional city in the USA, circumscribing together a city and its suburbs, sometimes known as an integrated labor market in economic geography. These definitions are updated regularly and available at `https://www.census.gov/programs-surveys/metro-micro/about/delineation-files.html`. The list of MSA included in analysis in the main text (Fig. 2.2) are enumerated in Table S1.

As the surveys from which we obtained depression prevalence estimates are administered by different agencies (the Substance Abuse and Mental Health Services Administration administers the NSDUH and the Centers for Disease Control and Prevention administers the BRFSS), collection and reporting methods differ substantially between these two data sources. The NSDUH is conducted in person while the BRFSS is conducted over the phone.

In addition, the two surveys differ in the questions they ask about depressive symptoms. The NSDUH asks participants whether or not they had a period of 2 or more weeks in which they experienced depressive symptoms in line with definitions in the DSM-IV Abuse and Administration [2011]. In contrast, the BRFSS asks respondents if they have ever been told that they "have a depressive disorder (including depression, major depression, dysthymia, or minor depression)?". In addition to these differences in questionnaire content and methods, the two data sources also differ in how they report data. The NSDUH reports age, ethnicity, and geography adjusted prevalence estimates in 33 metropolitan statistical areas (MSAs) Abuse and Administration [2011]. In contrast, the BRFSS reports age, gender, and socioeconomically adjusted prevalence estimates for any MSA with at least 500 respondents, and consequently the cities which are included in reports vary from year to year.

The 2010 NSDUH estimates of the rate of major depressive episodes used in Fig. 2.2 were obtained from Table 38 of the Substance and Mental Health Services Administration 2005 - 2010 National Survey on Drug Use and Health. This data are available online at `https://www.samhsa.gov/data/sites/default/files/NSDUHMetroBriefReports/NSDUHMetroBriefReports/NSDUH_Metro_Tables.pdf`. We multiplied estimated prevalence by 2010 estimated population to determine the estimate of total depression cases within each MSA.

The 2011-2017 BRFSS city estimates of the prevalence of major depression used in Fig. 2.2 are available online at `https://www.cdc.gov/brfss/smart/Smart_data.htm`. As with the NSDUH data we multiplied estimated prevalence by that years estimated population to estimate of total depression cases within each MSA.

One point of concern was that the cutoff of 500 respondents per city in the BRFSS data might artificially alter the joint distribution of prevalence and city size in a way that biases the estimate of β. One possibility is that larger cities are simply more likely to record enough responses to be included. However, since the BRFFS data includes cities with populations

as small as 20,285, whatever bias this 500 respondent cuttoff may introduce likely has a more complex origin. In order to address this without knowing the source of potential biases in city inclusion, we employed non-parametric change point detection based on the minimum covariate discriminant (MCD) Hubert and Debruyne [2010] in order to find the city size at which the joint distribution of city size and depression prevalence was different on either side of the change point. This was applied to the BRFSS data annually; results are consistent across years with the mean change point of (692,557 people, sd = 268,004 people).

Specifically, we followed a procedure similar to Cabrieto et al. [2017]. For each year of BRFSS data we first ordered the data by population and then applied a python implementation of the MCD algorithm Pedregosa et al. [2011] with a sliding window. This resulted in a robust-to-outliers estimate of the mean within the window and the 2-by-2 robust covariance matrix between population and depression prevalence within the window. These two quantities allow for the estimation of the Mahalanobis distance between the robust mean and the data from the city which has the next smallest population to the smallest city included in the window (the left out city). These distances follow a chi-squared distribution with degrees of freedom equal to the size of the window. Consequently, we marked the left out city as a potential change point if the Mahalanobis distance was greater than the 97.5th percentile of the relevant chi-squared distribution. Finally, we calculated a moving average with window size five, of marked change points. We considered a specific city size to be a change point if the moving average of marked change points was greater than 0.5. This was repeated for MCD window sizes from 5 to 25 data points in increments of 2. Histograms of the detected change points over all window sizes are shown in Fig. S1. When applied to the Twitter19' dataset change points are observed primarily at the ends of the population range (Fig. S2). This is suggestive of finite edge effects rather than a systematic change in the joint distribution of depression rates and city size as in seen in the BRFSS data.

Next we used a k-means clustering implementation in python Pedregosa et al. [2011]

to split the detected change points into two separate clusters based on the observation that the histograms of change points over all bin sizes for most years are roughly bimodally distributed. We used the two cluster centers as the final change points for each year of BRFSS data resulting in a partition of the data into three sets. Scaling estimates for the largest cities are reported in the main text and Fig. 2.2 and Table 2.1. When pooling all BRFSS data across all cities and years we still find evidence that larger cities have lower depression rates than smaller cities $\beta = 0.926$ (95% CI = [0.903, 0.950]). Results are similar when β is estimated separately for each year of BRFSS data (Table S4). When pooling data from the other two partitions which contain smaller cities, we found no evidence that depression rates scale sub- or super-linearly with population $\beta = 0.996$ (95% CI = [.956,1.035]) (Fig. S3). Results were similar when estimating β for each year separately (Table S2). This lack of a city size effect for smaller cities in the BRFSS data may indicate that social network determinants of depression are overshadowed by other risk factor in smaller cities, but may also be specific to biases introduced by the way in which the data were collected and reported.

We further estimated the sensitivity of our β estimates among larger cities to variation in the change point. For each year of BRFSS data we varied the change point 100 times according to a normal distribution with a mean of the change point used for that year in the main text and a variance equal to the variance in change points across years. We found that the estimates of β in larger cities are robust to these variations in the choice of change point (Table S3).

The geolocated Twitter dataset used in Fig. 2.2 and Table 2.1 is available online at `http://www.cs.cmu.edu/~ark/GeoText/`. This dataset included 377,616 tweets from 9,475 users collected over a one week period in March of 2010 Eisenstein et al. [2010]. Latitude and Longitude coordinates for each tweet were converted to a county-level Geographic Identifier (GEOID) using the US Census Geocoder API provided by the United States Census Bureau available at `https://github.com/fitnr/censusgeocode`. If there was more than one

coordinate per user, we used the mode and in the case of a tie, we used the coordinate that appeared first in time. We then used delineation files provided by the US Office of Budget and Management to roll up county level data to MSAs.

The Twitter19' dataset used in Fig. 2.2 was collected via Twitters academic research full search api (`https://developer.twitter.com/en/solutions/academic-research`) and was deemed not human subjects research by the University of Chicago IRB (IRB20-2049) due to the fact that all data are publicly available. Tweets that had available location tags (longitude and latitude) within U.S. cities between June 1st and July 1st 2019 were collected. This included data from 572,208 users and 15,076,651 tweets. The query parameters for retrieving tweets are available online at `https://github.com/enlberman/depression_scaling`. We note that while Twitter's opt-in policy for geo-location data changed in 2015 to require explicit consent to share precise GPS coordinates Cuomo et al. [2021], we rely on provided coarse location data included with all geo-located tweets.

We processed tweets following Yazdavar et al. [2017], using standard text preprocessing (for example, deleting stop words) and processing steps specific to the twitter platform (for example, deleting "#" in the hashtags). Then, we used a previously determined lexicon of seed terms related to depression symptoms organized into nine topics based on the PHQ-9, to guide a Latent Dirichlet Allocation (LDA) model Yazdavar et al. [2017]. LDA allows for the discovery of underlying topics within collections of text data and has been utilized previously with short, semi-structured text sources (e.g. Hong and Davison [2010], Schertz et al. [2018]). This enabled us to find users who had topic cluster(s) related to nine PHQ-9 topics in their tweets over one week.

One point of concern was that individuals who have depressive symptoms may tweet differently from those who don't have them. Specifically, we worried that individuals with depressive symptoms would tweet less leading to less reliable estimates from these individuals. This was the case: in the 2010 Twitter dataset, individuals with depressive symptoms

tweeted 57.7 times on average while individuals without depressive symptoms tweeted 37.6 times on average (t-statistic=25.7, p=7e-141). In order to control for this we performed a logistic regression to predict the presence of depressive language in users tweets from their number of tweets over the 1 week collection period. We repeated this procedure excluding users who had fewer than a specified number of tweets for cutoffs from 0 tweets to 110 tweets. As demonstrated in Fig. S4, the logistic regression model achieves significance for the 2010 Twitter dataset when individuals with fewer than 92 tweets are included. This indicates that people with depressive symptoms tend to tweet less, but that among individuals who tweeted at least 92 times over the collection period, a logistic regression model cannot differentiate between individuals with- and without- depressive symptoms based on their number of tweets. Consequently, we excluded individuals with fewer than 92 tweets and then estimated depression prevalence as the proportion of users in each city whose tweets contained a non-zero signal for any of the PHQ-9 topics. In the Twitter19' dataset, users with depressive symptoms tweeted 45.2 times on average and those without depressive symptoms tweeted 41.6 times on average (t-statistic = 9.52, p = 2e-21). Since the quantities of text were similar in both groups (compared to the 20 tweet difference in the 2010 Twitter dataset) we used a cutoff of 15 tweets to ensure that the LDA algorithm had sufficient input. In addition, we excluded cities in which estimated depression rates were unrealistic at 0% or 100%.

In order to test the sensitivity of the results to the minimum tweet threshold we repeated the scaling analysis on the Twitter dataset with minimum tweet count cutoffs from 82 to 101 (Table S5) and found that estimates of β were robust to these changes in exclusion criteria.

Estimating the Scaling Exponent β

We performed OLS linear regression in order to calculate the scaling exponent β for depression cases. We verified that the residuals of the models in Table 2.1 are approximately normally distributed both with q-q plots of the residuals (Fig. S5) and the the Shapiro-Wilk

test of normality Razali et al. [2011] (Table S6). We also verified that the residuals are not correlated with city size (Fig. S6, Spearman-r minimum p-value = .44).

Conditioning on Race, Education, Income, and Population Change

We additionally assessed whether city size was associated with a decreased risk of depression after conditioning on race, education, income, and population change. To do so we ran logistic regressions with the R package lme4 Bates et al. [2007] on each year of the BRFSS data using the individual participant level survey responses. We did this only for the 41 cities considered in the primary analysis. We used the BRFSS provided categories for income, race, and education. Consequently, the income variable had 6 levels with a baseline of not reported or missing, followed by: less than \$15k, \$15-\$25k, \$25-\$35k, \$35-\$55k, and greater than \$50k. The education variable had 5 levels with a baseline of not reported followed by: no high-school, graduated high-school, attended college, and graduated college. The race variable had 4 levels with a baseline of White followed by: Black, Asian, and other/multi-racial. We additionally included the natural logarithm of the population of each respondent's city as a dependent variable. The independent variable indicated whether each respondent had ever been told they have depression. The model is defined as

$$\begin{aligned} logit\{y_i = 1\} = \beta_0 + \beta_1 log(population) + \\ \beta_2 income_i + \beta_3 education_i + \\ \beta_4 race_i + \beta_5 \Delta population/population \end{aligned} \tag{2.2}$$

Results are summarized in Table S7, which were created with the R stargazer package Hlavac [2015]. The maximal 95% confidence interval for the odds ratio of log-city-population was found by taking the union of 95% confidence interval across all years of data.

Generative Network Model for Depression In Cities

The starting point for the simulations of depression cases was the log-skew-normal degree distribution, which has been shown to match the degree distributions of cell phone based social networks in cities Schläpfer et al. [2014a] and theoretically is the result of cumulative exposures to semi-random interactions taking place throughout cities' infrastructure networks. The log-skew-normal-distribution has the density function

$$p(\ln(k)) = \frac{2}{\omega}\phi(\frac{\ln(k)-\zeta}{\omega})\Phi(\alpha(\frac{\ln(k)-\zeta}{\omega})) \tag{2.3}$$

where ζ is the location parameter, α is the shape parameter, and ω is the scale parameter, ϕ is the normal distribution probability density function, and Φ is the normal distribution cumulative density function. These parameters can be transformed into the more familiar mean (μ), variance (σ^2), and skewness (γ_1) via Arellano-Valle and Azzalini [2013]:

$$\mu = \zeta + \omega\delta\sqrt{\frac{2}{\pi}}, \qquad \text{where} \quad \delta = \frac{\alpha}{\sqrt{1+\alpha^2}} \tag{2.4}$$

$$\sigma^2 = \omega^2(1 - \frac{2\delta^2}{\pi}), \tag{2.5}$$

$$\gamma_1 = \frac{4-\pi}{2}\frac{(\delta\sqrt{2/\pi})^3}{(1-2\delta^2/\pi)^{3/2}}. \tag{2.6}$$

We started with values of $\sigma = .87$, $\gamma_1 = .2$, and $\mu = 1.97$ in line with a city of size $N = 10,000$ Schläpfer et al. [2014a]. We then let the mean of this distribution grow with population size according to

$$\mu(N) = 1.97 + \delta \cdot \ln(\frac{N}{10,000}), \qquad \text{where} \quad \delta = \frac{1}{6} \simeq .167 \tag{2.7}$$

so that $\langle k \rangle \sim N^\delta$. For each simulated city with size, N, we sampled uniformly from it on a log scale from 10^4 to 10^7. We then sampled from the degree distribution N times to obtain

a list of the social network degrees of all N simulated city inhabitants. From this list we randomly assigned each simulated individual to be diagnosed with depression (or not) with a probability inversely proportional to their degree (probability of depression $\sim 1/\mathrm{k}$). Total depression cases in each simulated city were calculated as the sum of depressed individuals.

Code Availability

All relevant data processing code is available at `https://github.com/enlberman/depression_scaling`.

CHAPTER 3

EFFECTS OF RACIAL SEGREGATION ON ECONOMIC PRODUCTIVITY IN U.S. CITIES[1]

Homophily and heterophobia, the tendency for people with similar characteristics to preferentially interact with (or avoid) each other are pervasive in human social networks. Here, we develop an extension of the mathematical theory of urban scaling which describes the effects of homophily and heterophobia on social interactions and resulting economic outputs of cities. Empirical tests of our model show that increased residential racial heterophobia and segregation in U.S. cities are associated with reduced economic outputs and that the strength of this relationship increased throughout the 2010s. Our findings provide the means for the formal incorporation of general homophilic and heterophobic effects into theories of modern urban science and suggest that racial segregation is increasingly and adversely impacting the economic performance and connectivity of urban societies in the U.S.

Introduction

Homophily and heterophobia, the tendency of more similar individuals to preferentially interact and avoid interactions with others, are intimately familiar: most real-world McPherson et al. [2001] and digital Thelwall [2009] social networks show some degree of increased connectivity within certain groups and decreased connectivity between groups. Whether these preferences occur across characteristics of morality Dehghani et al. [2016], race Mollica et al. [2003], or nationality Thelwall [2009], minor individual preferences amplified by structural

1. This chapter has been posted as a preprint: Stier, Andrew J., Sina Sajjadi, Lus Bettencourt, Fariba Karimi, and Marc G. Berman. "Effects of Racial Segregation on Economic Productivity in US Cities." arXiv preprint arXiv:2212.03147 (2022).

proximity can result in large group-level differences Kossinets and Watts [2009], Schelling [1971, 2006], Dall'Asta et al. [2008]. Moreover, observed group differences in connectivity (outcome homophily/heterophobia) can have tremendous impacts on human behavior Karimi et al. [2018], Lee et al. [2019] despite not always being the result of strong individual choice preferences. In general, these effects expose individuals to less variety, knowledge and choice so that they slow down learning Golub and Jackson [2012], increase cognitive biases Lee et al. [2019], and limit the spread of information Halberstam and Knight [2016] among other detrimental effects Ertug et al. [2022], Ibarra [1992].

In cities, these effects play out across various types of preferences: people tend to travel between neighborhoods with similar socioeconomic demographics Heine et al. [2021], patterns of violent crime Graif et al. [2017a], and overall well-being Lathia et al. [2012a]. There is some evidence that spatial proximity in central locations can combat homophily and heterophobia, suggesting that these effects play out over large distances and more peripheral settings in cities Xu et al. [2019]. In addition, long histories of racism have led to spatial segregation among racial groups, particularly in the United States (U.S.) Kruse [2013], Nardone et al. [2020]. However, despite the universality of these effects in urban environments and their many pernicious effects, homophily and heterophobia have not yet been formally incorporated into the theoretical framework of modern urban science Bettencourt [2013, 2021b], which often assumes homogeneous (non-homophilic/heterophobic) mixing. Here we begin this process by developing homophily and heterophobia adjustments to the equations of urban scaling theory. We validate these adjustments empirically and provide evidence that racial heterophobia at the city level is predictive of lower overall economic productivity in the U.S. and that the strength of this relationship increased throughout the 2010's.

Results

Homophily and Heterophobia in Urban Scaling Theory

Urban scaling theory Bettencourt [2013, 2021b], which provides a theoretical backbone of modern urban science, describes cities as spatially embedded networks of socioeconomic interactions. Cities arise when the benefits of agglomerative increases in socio-economic outputs (denser socioeconomic networks) are balanced with the costs of maintaining infrastructure networks and transporting goods, services, and individuals throughout a city. These considerations result, under population averaging, in urban scaling laws that describe how different urban quantities scale with city size, N, defined as the number of individuals living in a city Bettencourt [2013, 2021b]. These empirically validated scaling laws have been found to hold for many cities across cultures and human history Bettencourt [2013, 2021b], Ortman et al. [2014a, 2016, 2015].

For the average per-capita number of social interactions, k, the urban scaling law takes form of $k \sim N^{\delta}$, where $\delta = \frac{1}{6}$. Corrections to these exponent values due to growth rate fluctuations and other higher-order effects, provide the basis for a statistical theory of urban scaling Bettencourt [2021b, 2020]. Nevertheless, the simplest and most widely used form of this scaling law results from a mean-field approximation that individuals within a city interact with others homogeneously, without restrictions of group affiliation. Under such conditions, taking individuals to have an interaction cross section a_0 and a characteristic travel length l per unit time, this approximation gives the average number of interactions for a large city ($N >> 1$) as given by:

$$k \sim \frac{a_0 l}{A} N \tag{3.1}$$

, where A is the area of the city's networks.

The scaling law for k can be recovered from Equation 3.1 by substituting the scaling

law for the area of the city's networks, $A \sim N^{1-\delta}$ Bettencourt [2013], which is the result of self-consistently balancing the net benefits of socioeconomic interactions with costs of transportation (and housing) overbuilt urban spaces.

Importantly, $\frac{a_0 l}{A}$ takes the role of a probability built out of the fraction of a city's area that individuals cover on average over a given time period. This is the average probability of interacting with all other individuals in the city, N. Thus, the total expected number of interactions for each individual (Equation 3.1) is given by their probability of interacting, $\frac{a_0 l}{A}$, multiplied by the number of individuals they could interact with, N.

Equation 3.1, assumes that all individuals in the city are equally likely to interact Bettencourt [2013]. However, this assumption is unrealistic and can be relaxed by assuming that individuals belong to a number of distinct groups, which in turn have group-specific interaction rates.

More specifically we model individuals in each of these groups as interacting preferentially with others of the same group, and with a lower probability with other groups. We define this relative reduction in out-group interactions by $1-h_g^{het}$, where $h_g^{het} \in [0,1]$ is the heterophobia of group g. Thus, $h_g^{het} = 0.8$ means that individuals from group g will only interact with 20% of out-group members, on average. However, when individuals lose contact with other groups, they may compensate by having a higher rate of intra-group interactions, depending on the social context Mollenhorst et al. [2008], Skvoretz [2013]. We model this compensatory effect with a similar parameter, $h_g^{hom} \in [0,1]$, which is the homophily of group g and sets the relative rate of intra-group interactions at $1+h_g^{hom}$. Though, h_g^{het}, and h_g^{hom} are uncorrelated here by assumption, different social contexts may induce positive or negative correlations between h_g^{het}, and h_g^{hom} Mollenhorst et al. [2008], Skvoretz [2013]. In addition, individuals in a city have a limit on the number of interactions they can take part in during any given time period so that e.g., when all of an individual's interactions are within their own group they do not interact with other groups (complete outcome homophily and heterophobia).

Taking this into account, future work might specify a budget for interactions that gives individuals the flexibility to trade between large numbers of less costly interactions (e.g., accessible within-group interactions) and fewer numbers of more costly but more rewarding interactions (e.g., to increase diversity and achieve superior group-level problem solving abilities Barkoczi and Galesic [2016]).

It is important to note that here, h_g^{hom} and h_g^{het} are understood to represent outcome homophily and heterophobia and do not prescribe strong individual choices (i.e., strong individual preferences or avoidance of groups). In other words, h_g^{hom} and h_g^{het} are the result of observed network segregation of group g that results from a combination of structural segregation and, possibly, individual preferences for certain groups. Though we expect relative increases in the rates of within-group interactions and relative decreases in the rate of between-group interactions, our model is agnostic to the direction of these effects (see Supplementary Text).

With these definitions, the average number of interactions for individuals in group g is given by:

$$k_g \sim \frac{a_0 l}{A}[N_g \cdot (1 + h_g^{hom}) + \sum_{j \neq g} N_j (1 - h_g^{het})] \tag{3.2}$$

where N_g is the population of focal group g. The total number of social interactions for all individuals in group g is $k_g N_g$, on average. Therefore, the average number of social interactions for individuals in a observed segregated city i with G different groups, is $k_i \sim \frac{1}{N} \sum_{g=1}^{G} k_{g,i} N_{g,i}$. Here the average number of interactions for each group, $k_{g,i}$ and the size of each group, $N_{g,i}$ are specific to the observed city. This simplifies to (see Supplementary Text):

$$k_i \sim k_0 N_i^{\delta} \cdot (1 - A_i^{het} + A_i^{hom}) \cdot e^{\xi_i} \tag{3.3}$$

with

$$A_i^{het} = \sum_{g=1}^{G} \sum_{j=g+1}^{G} \frac{N_{g,i}}{N_i} \frac{N_{j,i}}{N_i} (h_{g,i}^{het} + h_{j,i}^{het}); \qquad A_i^{hom} = \sum_{g=1}^{G} (\frac{N_{g,i}}{N_i})^2 h_{g,i}^{hom}. \tag{3.4}$$

Here, k_0 is the scaling prefactor, A_i^{het} is the heterophobia adjustment, A_i^{hom} is the homophily adjustment, and ξ_i are additional city specific effects Bettencourt et al. [2010]. The A_i^{het}, A_i^{hom} are simply the averages of the coefficients $h_{g,i}^{het}, h_{g,i}^{hom}$, weighted by group sizes in each city. We see that $1 - A_i^{het} + A_i^{hom}$ gives a city's specific multiplicative adjustment to the scaling law. Note that $A_i^{het} \in [0, .25]$, while $A_i^{hom} \in [0, 1]$ for a city with at least two groups so that, in this realization of the model, interactions are increased within-groups and decreased between-groups (see Supplementary Text).

Thus, we expect that increased segregation between groups reduces social interactions (unless fully compensated for by increased within-group interactions), in line with previous research Ibarra [1992]. In addition, A_i^{het} depends on the relative sizes of groups and has the largest impact when all groups are of equal size (see Supplementary Text), matching previous investigations of homophily and heterophobia in small networks Oliveira et al. [2022].

The final step is to consider how interactions between city inhabitants translate to economic outcomes. In the derivation of urban scaling laws for economic outputs, interactions over various types (friendship, employment, acquaintance, etc) can couple to economic outputs either positively or negatively, and with varying strengths over the different types Bettencourt [2013]. Similarly, interactions between- or within-groups can couple deferentially to social and economic outputs, so that, for instance, within-group interactions might be more productive for social outputs, but less productive for creative outputs Barkoczi and Galesic [2016]. For simplicity, here, we assume that between and within group interactions do not couple to economic outputs differently, so that economic outputs are directly proportional to the social interactions specified by Equation 3.3 Bettencourt [2013, 2021b].

Empirical tests of the homophily and heterophobia adjustments

We next sought to test the empirical validity of the homophily and heterophobia adjustments regarding self-reported race in U.S. cities. A_i^{het} and A_i^{hom} were calculated from racial demo-

graphic estimates in cities collected by the U.S. census for each year between 2010-2020 (see Materials and Methods). Racial segregation for each city and group was calculated as the

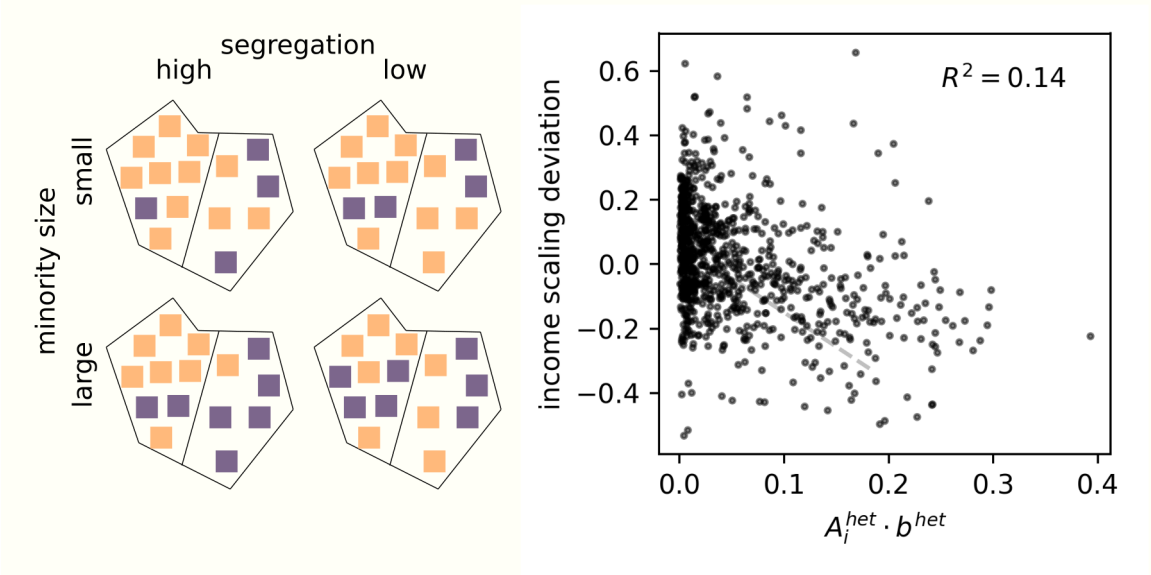


Figure 3.1: Left: Measures of citywide segregation depend on the relative size of the groups and their spatial concentration. Starting at the top left and going clockwise, $A_i^{het} \sim \sum \frac{N_{g,i}}{N_i} \frac{N_{j,i}}{N_i}(s_{g,i} + s_{j,i})$ takes on values of 0.92, 0.98, 0.98, and, 0.85. Here, $s_{g,i}$ is the level of residential segregation of group g in city i. The large minority, high segregation city in the lower left has the smallest A_i^{het}. Top: a city where the majority makes up 73% of the population and the average segregation is 20% (left) or 5.5% (right). Bottom: a city where the majority makes up 53% of the population and the average segregation is 30% (left) or 3% (right). Right: The relationship between urban scaling law deviations for median income and $A_i^{het} \cdot b^{het}$ in U.S. cities in 2020.

average difference between neighborhood proportions of residents in each racial group and the city-wide mean. Explicitly, $s_{g,i} = \frac{1}{M} \sum_{m=1}^{M} |N_{g,m,i}/N_{m,i} - N_{g,i}/N_i|$, where m indexes all of the neighborhoods in a city and $N_{g,m,i}/N_{m,i}$ is the proportion of residents in racial group g for neighborhood m in city i (see Figure 3.1; we note that residential segregation may only capture some of the overall social contact segregation that occurs in cities Tammaru et al. [2016], Priest et al. [2014], Tucker et al. [2021]). For example, in a city with two groups where the majority is 80% of the population, one neighborhood might have 90% of its population from the majority group while another neighborhood is 20% majority group; in this case, the

majority group would have a segregation value of $0.5 \cdot (|0.9 - 0.8| + |0.2 - 0.8|) = 0.35$. Since we expect increased segregation, on average, to lead to increased homophily and heterophobia values, though possibly with different strengths, we modeled these values as linearly related to the empirical segregation values: $h_i^{\cdot} \sim b^{\cdot} * s_i$, where $b^{\cdot}$ determines the strength of coupling between residential racial segregation and heterophobia. In order to ensure that our results were not sensitive to the choice of segregation measure we repeated analyses with four additional segregation measures (see Materials and Methods).

We performed these analyses for two measures of economic productivity, median income and gross domestic product (GDP), in order to assess the impacts of segregation on individual and overall economic productivity in U.S. cities, respectively. Though median income and GDP are correlated (Spearman r $\sim$ 0.55, see Supplementary Table 7.11), the ability of individuals to garner higher wages and of businesses to generate high economic outputs are not commensurate Nolan et al. [2019].

We chose to conduct our analyses at the level of census tracts for the U.S., which are small spatial units with approximately 4,000 inhabitants (results were similar when smaller spatial units of census block groups were used, see Supplementary Text). Analyses were conducted within functional cities (integrated commuting areas), defined as combined statistical areas by the U.S. Office of Budget and Management of Management and Budget [2021]. These are functional definitions of cities that capture where people live, socialize, and work Bettencourt [2021b], Stier et al. [2022d].

These analyses revealed the variation in scaling deviations explained by homophily and heterophobia and the empirical strength, $b^{het,hom}$, of A_i^{het} and A_i^{hom}, i.e., the degree to which segregation impacts interactions between and within groups, respectively. The results demonstrate that scaling deviations for income are significantly predicted by A_i^{het} across all years (Figure 3.1, Supplementary Table 7.12), but not by A_i^{hom} (Supplementary Table 7.12). Though A_i^{het} and A_i^{hom} are strongly correlated ($\bar{r} \sim 0.77$, range $[0.75, 0.78]$ across years),

the variance inflation factor is relatively low when using centered versions of these variables ($VIF < 2.58$ for all years) and the Akaike information criterion (AIC) of models with only A_i^{het} is always lower than models with only A_i^{hom} (average $\Delta\text{AIC} = -44$, range $[-54, -37]$).

In addition, A_i^{het} explains a maximum of 13.7% of the variance in scaling deviations for income in 2020 (Figure 3.1, Supplementary Table 7.13). In contrast, A_i^{het} is only significantly predictive of GDP scaling deviations after 2014 where it explains a maximum 1.3% of the variance in 2020 (Supplementary Table 7.14. Results were similar when alternate measures of residential racial segregation were used (see Materials and Methods, Supplementary Tables 7.15-7.20). These results suggest that heterophobia has a stronger effect on economic productivity than homophily and that it is more important for individual outcomes than for the whole of cities' economies.

One reason for these differences might be that the types of interactions that racial segregation curtails are more important for labor opportunities and associated opportunities for securing higher wages and incomes than they are for the overall productivity of firms in an urban economy Alabdulkareem et al. [2018]. These differences could be operationalized in future work through differential coupling of various modes of between- and within-group interactions to various types of economic outputs at different scales of organization.

Finally, we observe that the degree to which residential segregation is associated with economic productivity increased between 2010 and 2020 (Figure 3.2). In particular, for both median income and GDP, the explained variance of scaling deviations, R^2, increased systematically between 2010 and 2020. This happened while the average value of A_i^{het} stayed relatively constant (Supplementary Figure 7.1). Thus, though observed racial segregation in U.S. cities was relatively constant during the 2010s, the degree to which segregation accounted for lower incomes and GDP, relative to city size, likely increased.

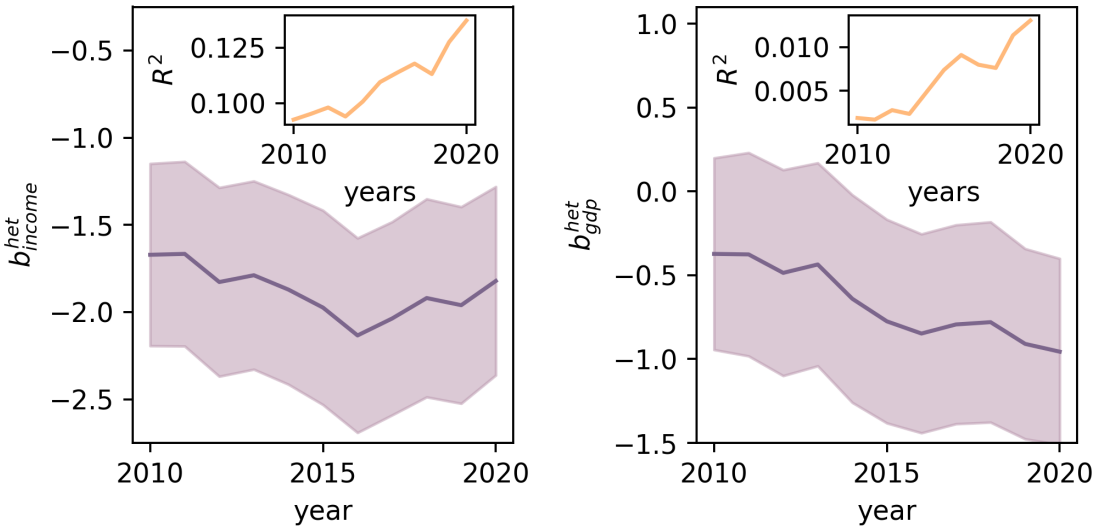


Figure 3.2: Changes in the empirical strength of the heterophobia adjustment over time. Insets show R^2 values for the OLS regression models. Shaded regions show the 95% confidence interval for b^{het}, i.e. the strength of the relationship between h_g^{het} and s_g. The different values of b^{het} are parameters in the OLS regression models. Left: median income. Right: gross domestic product (GDP).

Discussion

Our development of urban scaling relations to account for racial segregation effects suggests that heterogeneous network structure in cities can meaningfully impact their economic outputs. The adjustment to urban scaling laws we derived suggests that segregation reduces economic outputs with strength depending on the size of different groups and their levels of heterophobia Lee et al. [2019]. Our empirical findings support the hypothesis that this is indeed the case. For example, in 2020, the New Orleans, Louisiana Metropolitan Area had a median income of approximately $54,400. The urban scaling law predicts a median income of approximately $65,000 from city size alone, while accounting for segregation brings the prediction to approximately $56,500.

In general, interactions beyond the residence, in shared public spaces and in workplace environments, are also likely relevant to economic outputs. Consequently, it is important

that future work model heterogeneous, segregated ambient mixing in these environments and evaluate it alongside residential segregation. Such considerations are particularly important for understanding, more generally, how differences in structured group interactions lead to more or less productive cities.

Moreover, our observation that the influence of heterophobic preferences on economic outputs is increasing over time suggests that additional factors are accelerating the detrimental effects of segregation in United States cities. These might include factors like peer influence and behavioral norms Jackson [2021] which can interact with heterophobia to exacerbate induced inequalities.

Why do individual economic outcomes (e.g., income) show stronger associations with racial segregation than overall economic output? How does racial segregation effect other well-described urban social behaviors such as mental health Stier et al. [2021] and crime Bettencourt et al. [2010]? How do other forms of segregation of interactions, via e.g., politics and wealth, impact the various economic and social outputs? Such questions are important for future research and require further development of theoretical models and empirical investigations along the lines developed here. It is important that this research is performed, at least in part, in the context of cities, which while housing enormous human diversity often fail to make the most of their latent socioeconomic potential. Urban environments are now home to the majority of the world's population and account for a disproportionate fraction of global economic and intellectual productivity; as a result, a better understanding of the sources of inequality, and successful integration of diversity in cities is crucial to building more equitable and inclusive societies.

Materials and Methods

U.S. Census and Economic Data

All census data is publicly available and was downloaded from `data.census.gov`. Five year racial demographic estimates for U.S. census tracts and census block groups were downloaded from table *B02001*. Homophily values were calculated across the six racial groups provided in these tables: White, Black, Native American/Native Alaskan, Asian, Hawaiian/Pacific Islander, and Other. Five-year population estimates for U.S. cities defined as combined statistical areas (CBSAs) were downloaded from table *B01003*. We note that these demographic estimates are only available for block groups from 2013 onward, but are available for census tracts from 2010 onward. Five-year median income estimates for U.S. cities were downloaded from table *B19013*. Gross domestic product data for CBSAs is publicly available from the Bureau of Economic Analysis and was downloaded from `https://apps.bea.gov/iTable/index_regional.cfm`. 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`.

Association Between Scaling Deviations and A_i^{het} and A_i^{hom}

In order to determine the association between A_i^{het} and A_i^{hom} and measures of economic productivity, the scaling relationship between economic measures and city population has to be removed. However, we first recognize that our measured segregation values, though $\in [0, 1]$ are not true probabilities. However, we can expect $h_{g,i}^{hom}$ and $h_{g,i}^{het}$ to be proportional to the empirical segregation values. Explicitly,

$$h_{g,i}^{\cdot} = h^{\cdot} + b^{\cdot} s_{g,i} \tag{3.5}$$

, where $h^{\cdot}$ is the base probability, $b^{\cdot}$ a scaling factor, and $s_{g,i}$ the empirical segregation values for group g in city i. We have, for simplicity, assumed a linear relationship so that the hetrophobia and homophily values increase linearly with the degree of segregation.

A_i^{het} and A_i^{hom} then become:

$$A_i^{het} = \sum_{g=1}^{G} \sum_{j=g+1}^{G} \frac{N_{g,i}}{N_i} \frac{N_{j,i}}{N_i} [2 \cdot h^{het} + b^{het} \cdot (s_{g,i} + s_{g,j})]; \;\; A_i^{hom} = \sum_{g=1}^{G} (\frac{N_{g,i}}{N_i})^2 \cdot [h^{hom} + b^{hom} \cdot s_{g,i})] \tag{3.6}$$

when the combined homopohily and heterophobia effects are small we can approximate this by ($ln(1+x) \sim x$ when $x << 1$):

$$\begin{aligned} ln(k_i) \simeq \delta \cdot ln(N_i) - \sum_{g=1}^{G} \sum_{j=g+1}^{G} \frac{N_{g,i}}{N_i} \frac{N_{j,i}}{N_i} [2 \cdot h^{het} \\ + b^{het} \cdot (s_{g,i} + s_{g,j})] + \sum_{g=1}^{G} (\frac{N_{g,i}}{N_i})^2 \cdot [h^{hom} + b^{hom} \cdot s_{g,i}] + \xi_i \end{aligned} \tag{3.7}$$

since our primary interest is on the effects of segregation, we can write this as two regression equations:

$$ln(k_i) \sim ln(C) + \beta \cdot ln(N_i) + \epsilon_i \tag{3.8}$$

and,

$$\epsilon_i \sim D - b^{het} \cdot \sum_{g=1}^{G} \sum_{j=g+1}^{G} \frac{N_{g,i}}{N_i} \frac{N_{j,i}}{N_i} (s_{g,i} + s_{j,i}) + b^{hom} \cdot \sum_{g=1}^{G} (\frac{N_{g,i}}{N_i})^2 \cdot s_{g,i} + \xi_i \tag{3.9}$$

where we have included the additional city specific terms of Equation 3.7 in the residuals ξ_i which are the same city specific effects from Equation 3.3. Here, C and D are city size independent constants, i.e., scaling prefactors. In addition, β is the scaling exponent estimate which we expect to take on the value of $\delta = \frac{1}{6}$.

We estimated the regression for Equation 3.8 by OLS first and then used the scaling

deviations, ϵ_i, from those regressions to estimate the influence of homophily and heterophobia by OLS via Equation 3.9. In order to exclude outlier cities that significantly deviate from the scaling law, cities with $|\epsilon| > 3.09\sqrt{Var(\epsilon)}$, i.e., beyond the 99.9%th percentile of the normal distribution of the standard deviation of ϵ were excluded for each year. Results are similar when outliers are not excluded (Supplementary Tables 7.21 & 7.22).

Alternate Measures of Residential Segregation

In order to ensure that the results were not sensitive to a specific segregation measure we repeated analyses with three additional measures of residential segregation White [1986]. These included the normalized segregation index:

$$D_{g,i} = \frac{\sum_m |\frac{N_{g,m,i}}{N_{m,i}} - \frac{N_{g,i}}{N_i}| \cdot N_{m,i}}{2 \cdot N_i \cdot \frac{N_{g,i}}{N_i} \cdot (1 - \frac{N_{g,i}}{N_i})} \tag{3.10}$$

the Gini Coefficient:

$$gini_{g,i} = \frac{\sum_m \sum_l |\frac{N_{g,m,i}}{N_{m,i}} - \frac{N_{g,l,i}}{N_{l,i}}| \cdot N_{m,i} \cdot N_{l,i}}{2 \cdot N_i^2 \cdot \frac{N_{g,i}}{N_i} \cdot (1 - \frac{N_{g,i}}{N_i})} \tag{3.11}$$

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

$$\eta^2_{g,i} = \frac{\sum_m N^2_{g,m,i}}{N_{g,i} \cdot (1 - \frac{N_{g,i}}{N_i})} - \frac{\frac{N_{g,i}}{N_i}}{1 - \frac{N_{g,i}}{N_i}} \tag{3.12}$$

The construction of measures of segregation based more directly on interactions, beyond the composition of residential neighborhoods, is also likely important and will be pursued in future work.

Data Availability

All data used to support the results were downloaded from publicly available government sources. Residential demographics were downloaded via data.census.gov from table B02001. Population estimates were downloaded from table B01003. Median income estimates were downloaded from table B19013. Gross Domestic Product data was downloaded from the Bureau of economic analysis via https://apps.bea.gov/iTable/index_regional.cfm. Delineation files to map between census tracts, blocks, and CBSAs were downloaded from the U.S. Office of Budget and Management from https://www.census.gov/programs-surveys/metro-micro/about/delineation-files.html

Code Availability

All analyses and figures were made with python code, using standard python libraries for scientific computing including statsmodels, matplotlib, numpy, and scipy. All code necessary to reproduce the results is available at https://github.com/enlberman/racial_segregation_economic_productivity

Competing Interests

The authors declare no competing interests.

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 or are publicly available, the Supplementary Materials, or are publicly available.

Author contributions: A.J.S. and F.K. 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.

CHAPTER 4

CITY POPULATION, MAJORITY GROUP SIZE, AND RESIDENTIAL SEGREGATION DRIVE IMPLICIT RACIAL BIASES IN U.S. CITIES[1]

Implicit biases - differential attitudes towards out-group members - are pervasive in human societies. These biases are often racial in nature and create inequities across many aspects of life. Recent research has revealed that implicit biases are, generally, driven by social contexts. However, it is unclear if the regular ways that humans self-organize in cities systematically influence implicit racial bias strength. We leverage extensions of the models of urban scaling theory to predict and test between-city differences in these biases. Our model links spatial scales from city-wide infrastructure to individual psychology to predict that cities that are more populous, more diverse, and less segregated are less biased. We find broad empirical support for these predictions in U.S. cities with data spanning a decade from millions of individuals. We conclude that the organization of cities strongly drives the strength of these biases and provides potential systematic intervention targets for planning more equitable societies.

Introduction

Cities are organized in surprisingly regular ways Bettencourt [2013, 2021b], Molinero and Thurner [2021], which drive and constrain social interactions similarly across cultures and time Schläpfer et al. [2014b], Oliveira et al. [2017], Lobo et al. [2020]. However, there are

1. This chapter is posted as a preprint as: Stier, Andrew, Sina Sajjadi, Fariba Karimi, Luis Bettencourt, and Marc G. Berman. "City Population, Majority Group Size, and Residential Segregation Drive Implicit Racial Biases in US Cities." Majority Group Size, and Residential Segregation Drive Implicit Racial Biases in US Cities (January 27, 2023) (2023).

many factors beyond the built-space geometry Bettencourt [2021b], Molinero and Thurner [2021] of cities that modulate urban social interactions. Among these, implicit biases towards out-group members are one of the most universal Dunham et al. [2006]. 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 Dehon et al. [2017], scholastic performance Jacoby-Senghor et al. [2016], employment Ziegert and Hanges [2005], policing Ekstrom et al. [2022], Hehman et al. [2018], mental health outcomes Chae et al. [2017], and physical health Chae et al. [2012]. 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 Baron and Banaji [2006], Gibson et al. [2017], are stable into adulthood, and are less prevalent in schools with more diverse populations Gibson et al. [2017]. 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 Kubota et al. [2012], Amodio [2014], Cloutier et al. [2017], Brosch et al. [2013]. 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 Payne and Vuletich [2018], Jackson et al. [2014], Boscardin [2015].

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 Gonzalez et al. [2017], Lai et al. [2014, 2016]. This suggests that individuals' biases likely reflect ongoing predictions about their social environment Payne and Hannay [2021], De Houwer [2019], and consequently, that consistent population averages of implicit biases Vuletich and Payne [2019] 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 Payne et al. [2019], Payne and Hannay [2021], De Houwer [2019], Vuletich and Payne [2019].

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 Payne et al. [2019]. 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 Bettencourt et al. [2010], Bettencourt [2013, 2021b], 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 Bettencourt [2013, 2021b]. 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 Bettencourt [2013, 2021b]. These models derive average properties

of cities as a function of their population size, N, as scale-invariant scaling laws Bettencourt [2013, 2021b]. For example, in the case of average per-capita social interactions, k, the scaling law takes the form of $k \sim N^{\delta}$, where $\delta = \frac{1}{6}$.

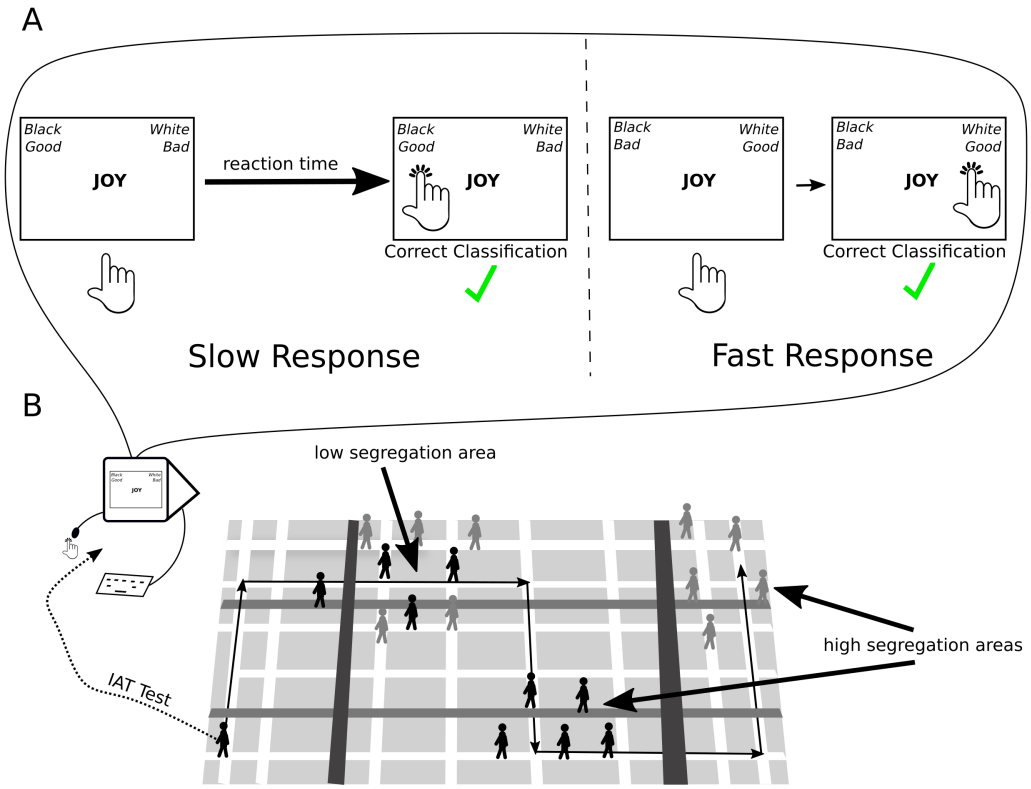


Figure 4.1: A) The Implicit Association Test measures implicit racial biases as a relative difference in reaction times between different pairings of word and face categories. B) We model implicit racial biases in cities as a cumulative exposure process to out-group individuals shaped by city population 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 Stier et al. [2022c]. 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:

$$k_i \sim N_i^{\delta} \left[\sum_{g=1}^{G} (\frac{N_{g,i}}{N_i})^2 (1 + h_{g,i}^{hom}) + \sum_{g=1}^{G} \sum_{j \neq g} \frac{N_{g,i}}{N_i} \frac{N_{j,i}}{N_i} (1 - h_{g,i}^{het}) \right] \tag{4.1}$$

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 $1 + h_{g,i}^{hom}$ Stier et al. [2022c]. 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 Stier et al. [2022c].

The first term of Equation 4.1 is the typical scaling law Bettencourt [2013, 2021b], Schläpfer et al. [2014b]. 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 Gibson et al. [2017], Gonzalez et al. [2017], Payne and Hannay [2021], De Houwer [2019], Laurence [2011], Pettigrew et al. [2007], Allport et al. [1954], Wagner et al. [2003, 2006], 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 Bettencourt [2013]. Previous research has suggested that this relationship is positive – more inter-group interactions are associated with lower implicit bias levels Gibson et al. [2017], Gonzalez et al. [2017], Payne and Hannay [2021], De Houwer [2019], Laurence [2011], Pettigrew et al. [2007], Allport et al. [1954], Wagner et al. [2003, 2006]. 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 Kubota et al. [2012], Amodio [2014], Cloutier et al. [2017], Brosch et al. [2013].

A common feature of such expertise-based learning is decreasing marginal returns to exposure, which is often formalized in a learning curve Crossman [1959], Woźniak et al. [1995], Murre and Dros [2015], Van der Zwaan and Rabl [2003]. Learning curves describe the relationship between costs and expertise across diverse individual or group tasks such as motor learning Woźniak et al. [1995], sequence learning Murre and Dros [2015], solar panel construction Van der Zwaan and Rabl [2003], and cigar rolling Crossman [1959]. 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) Bettencourt [2013, 2021b], 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:

$$b_i \sim N_i^{-\delta\alpha} \cdot \left[\sum_{g=1}^{G} \sum_{j \neq g} \frac{N_{g,i}}{N_i} \frac{N_{j,i}}{N_i} (1 - h_{g,i}^{het}) \right]^{-\alpha} \tag{4.2}$$

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):

$$b_i \sim N_i^{-\delta\alpha} \cdot \left[\frac{N_{1,i}}{N_i} - (\frac{N_{1,i}}{N_i})^2 \right]^{-\alpha} \cdot (2 - h_{1,i}^{het} - h_{2,i}^{het})^{-\alpha} \tag{4.3}$$

Equation 4.3 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 7.1), 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 4.3 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 4.3. 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 Cesaretti et al. [2016], growth rate fluctuations, and other higher-order effects Bettencourt [2020], which may contribute to differences in independent estimates of α calculated from the first and the second terms of Equation 4.3.

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 Xu et al. [2014]. 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 Stier et al. [2022d]. 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 Stier et al. [2022c]. 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 4.3 (Figure 4.2A, Supplementary Table 7.23).

In addition, larger majority group sizes and higher levels of residential racial segregation are significantly related to scaling deviations (Supplementary Table 7.23) and associated with higher average IAT scores, in line with Equation 4.3. We note that when analyzing

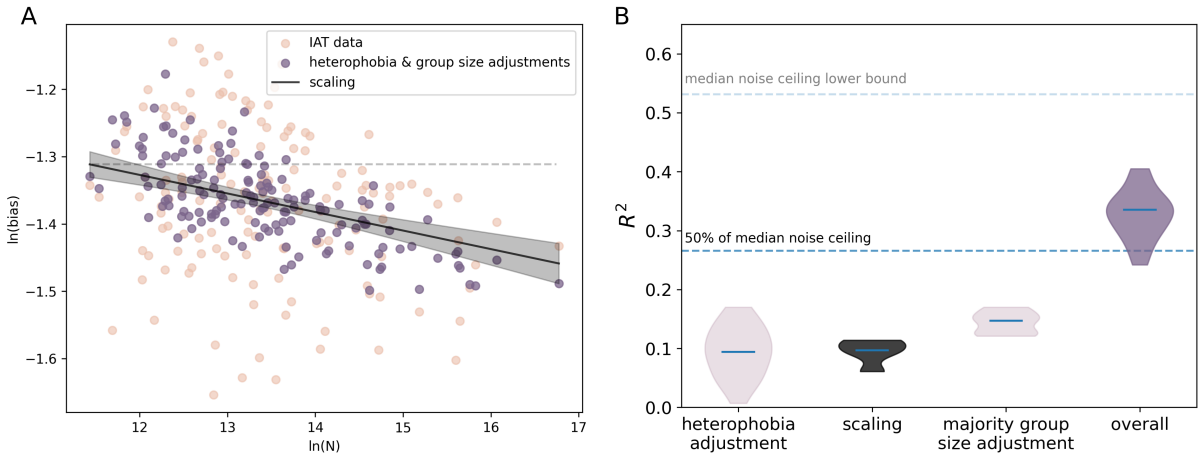


Figure 4.2: A) Scaling relationship, majority group size adjustment, and heterophobia adjustment for IAT data from 2020 in 149 cities with > 500 IAT responses per city. The shaded region is the 95% confidence interval for the scaling relationship. For visualization purposes, the heterophobia adjustment 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 7.59-7.66). 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.

single years of 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 7.24).

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 7.25-7.35). 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 Hehman et al. [2021]. 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 7.36-7.39). 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 Tucker et al. [2021], 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 4.2B, Supplementary Tables 7.40-7.46). Estimates of the noise ceiling Storrs et al. [2020] suggest that these three structural factors may capture a majority of the variance than can be accounted for given the reliability Schnabel et al. [2008] of the IAT measure (Supplementary Tables 7.47-7.58). 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 4.2B).

Finally, we compared estimates of the learning rate, α to previously conducted experimental interventions Lai et al. [2014, 2016] 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 4.3). 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 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 Payne et al. [2017], Lee et al. [2019]. Importantly, our model suggests that implicit racial biases emerge from the interaction between city-wide social contexts that are shaped

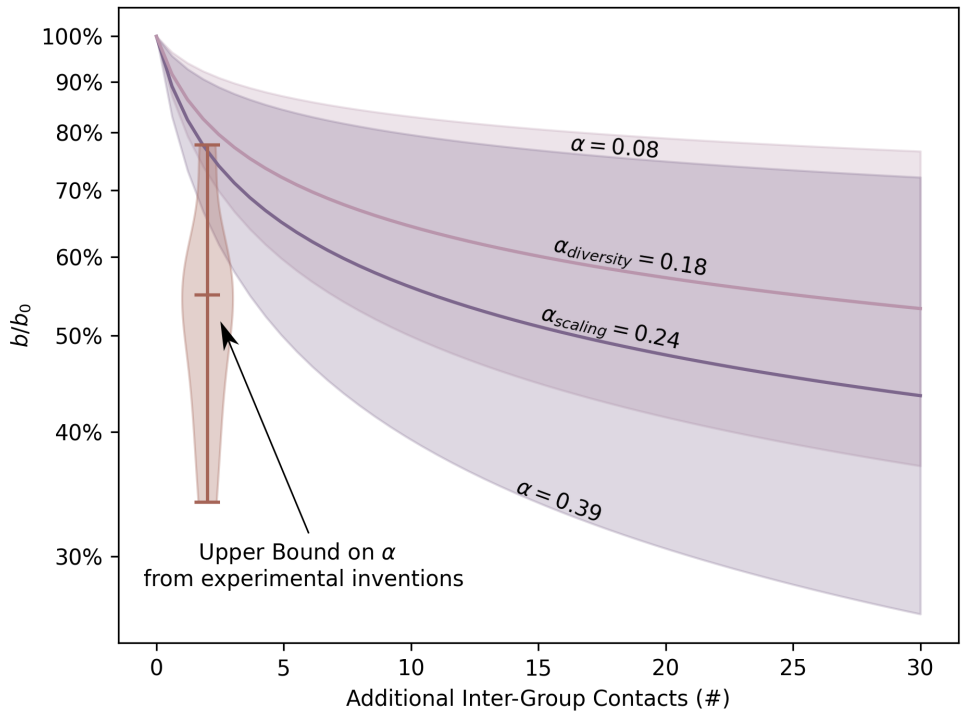


Figure 4.3: Estimated learning rates, α. We plot learning as a decrease in bias levels relative to an arbitrary baseline, $\frac{b}{b_0}$ as a function of the number of additional inter-group contacts. Solid curves indicate the mean estimated learning rate from the scaling exponent or majority group 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 Lai et al. [2014, 2016] designed to simulate one-shot inter-group contact of varying quality.

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 Greenwald et al. [2015]. This is important to recognize as cities with lower levels of racial segregation also tend to, not accidentally, have higher incomes Stier et al. [2022c] and healthier inhabitants Kramer and Hogue [2009].

These results, along with our related work Stier et al. [2022c] 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 Bak-Coleman et al. [2021], Dalege et al. [2017] and issues of trust and collective action, for example relating to public health programs such as vaccines Galesic et al. [2021], Dalege and van der Does [2022].

Materials and Methods

IAT Data

All racial IAT Data are publicly available Xu et al. [2014] 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 Greenwald et al. [2003] 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 7.59-7.66).

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:

$$\ln(b_i) \sim C + \beta_1 \cdot \ln(N_i) + \epsilon_i \tag{4.4}$$

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:

$$\epsilon_i \sim C_2 + \beta_2 \cdot \ln(\frac{N_{1,i}}{N_i} - \frac{N_{1,i}^2}{N_i^2}) + \beta_3 \cdot \ln(2 - h_{1,i} - h_{2,i}) + \xi_i \tag{4.5}$$

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 Stier et al. [2022c] 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 4.5 then becomes:

$$\epsilon_i \simeq C_2 + \beta_2 \cdot \ln(\frac{N_{1,i}}{N_i} - \frac{N_{1,i}^2}{N_i^2}) - \frac{\beta_3}{2} \cdot [2 \cdot h^{het} + b^{het} \cdot (s_{1,i} + s_{2,i})] + \beta_3 \ln(2) + \xi_i \tag{4.6}$$

where we have substituted the heterophobia values via the equation $h_{g,i} = h^{het} + b^{het} s_{g,i}$ Stier et al. [2022c]. 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:

$$\epsilon_i \simeq C_2 + \beta_2 \cdot \ln(\frac{N_{1,i}}{N_i} - \frac{N_{1,i}^2}{N_i^2}) + \beta_3 \cdot (s_{1,i} + s_{2,i}) + \xi_i \tag{4.7}$$

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 Storrs et al. [2020].

Measures of Residential Segregation

As in our related work Stier et al. [2022c], all analyses were conducted across four different measures of residential segregation White [1986] in order to ensure that the results were not sensitive to any specific metric. These included the mean deviance measure:

$$\Delta_{g,i} = \frac{1}{M} \sum_{m}^{M} |N_{g,m,i}/N_{m,i} - N_{g,i}/N_i|, \tag{4.8}$$

the normalized segregation index:

$$D_{g,i} = \frac{\sum_m |\frac{N_{g,m,i}}{N_{m,i}} - \frac{N_{g,i}}{N_i}| \cdot N_{m,i}}{2 \cdot N_i \cdot \frac{N_{g,i}}{N_i} \cdot (1 - \frac{N_{g,i}}{N_i})}, \tag{4.9}$$

the Gini Coefficient:

$$gini_{g,i} = \frac{\sum_m \sum_l |\frac{N_{g,m,i}}{N_{m,i}} - \frac{N_{g,l,i}}{N_{l,i}}| \cdot N_{m,i} \cdot N_{l,i}}{2 \cdot N_i^2 \cdot \frac{N_{g,i}}{N_i} \cdot (1 - \frac{N_{g,i}}{N_i})}, \tag{4.10}$$

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

$$\eta^2_{g,i} = \frac{\sum_m N^2_{g,m,i}}{N_{g,i} \cdot (1 - \frac{N_{g,i}}{N_i})} - \frac{\frac{N_{g,i}}{N_i}}{1 - \frac{N_{g,i}}{N_i}}. \tag{4.11}$$

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 Hehman et al. [2021]. 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 Kind and Buckingham [2018], of Wisconsin School of Medicine and Health. [2019]. 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 Hehman et al. [2021].

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 4.5 and 4.7. From Equation 4.5 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 4.3 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 Lai et al. [2014], Lage-Castellanos et al. [2019]. 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 Lai et al. [2014, 2016]. 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.

Acknowledgments

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

The authors thank Shige Oishi and Nicholas Epley for their helpful discussions during the preparation of the manuscript.

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.

Data Availability

All data needed to evaluate the conclusions in the paper are present in the paper, the Supplementary Materials, or are publicly available.

Competing Interests

The authors declare that they have no competing interests.

CHAPTER 5

EVIDENCE FOR IMPROVED SELECTIVE ATTENTION IN LARGER U.S. CITES[1]

Attentional processes are fundamental to managing the many sources of information that are crucial to city life. Historically, these many demands on urban dwellers' lives were assumed to induce "failures of attention" associated with negative aspects of city life, including stress, mental fatigue, and low social responsibility. However, these studies have not, in general, directly measured attention, or they have applied coarse characterizations of attention. Here we directly measure selective attention via a smartphone based task in 3,082 individuals and compare performance across cities of different sizes. We find that individuals in more populous cities tend to have better selective attention performance. Further, the quantitative details of this observation suggested that this effect is fundamentally related to socioeconomic interactions in cities. Further, interpreting these results within the framework of Urban Scaling Theory suggests that differential attentional abilities between cities may have consequences for interaction quality but not interaction quantity. These results suggest that the assumed relationship between urbanicity needs to be reevaluated and provides new directions of research for psychology and urban science.

Introduction

Cities are full of distractions. At any given moment city dwellers are simultaneously managing social interactions, avoiding moving obstacles, navigating complex street networks, and

1. The co-authors for this chapter are: Jillian Rae Silva-Jones, Monica D. Rosenberg, Luís M.A. Bettencourt, Lauren N. Whitehurst, and Marc G. Berman.

filtering relevant sounds from the urban din. The ability to attend to each of these stimuli, in turn, and when necessary, is fundamental to city life.

At the turn of the 20th century, the constant demands on attention in urban contexts became a cause for concern. Sociologists suggested that due to the large numbers of social contacts that people have in cities, they necessarily have larger quantities of superficial relationships than individuals outside of cities Simmel [1950], Wirth [1938]. Similarly, psychologists started to document the cognitive adaptations associated with cities Milgram [1970a], including a faster pace of life (e.g., walking speed Bettencourt et al. [2007a] and the brusqueness of interactions Milgram [1970a]), and sparse cognitive maps of cities (i.e., familiarity with only areas of cities that a relevant to each individual Milgram [1970a]). For the most part, these early works implicitly focused on how individuals selectively allocate finite attentional resources to various urban stimuli. These "failures of attention" were, in turn, associated with "stimulus overload" that can induce negative aspects of city life White and Shah [2019], for example, stress (6–9), mental fatigue (10), low social responsibility Milgram [1970a], and, greater psychological impacts of crime Milgram [1970a].

As a result of these observations, a common assumption has been that the "overload" characteristics of cities lead to poorer attentional capacity Wilson [1986], Ulrich [1993], Kaplan [1995], Berman et al. [2008]. For example, the biophilia Wilson [1986], Ulrich [1993] and attention restoration Kaplan [1995], Berman et al. [2008] hypotheses imply that the lack of natural features in urban environments results in fatigued attention. Further, since attention supports cognition generally Fisher [2019], Burgoyne and Engle [2020], the assumption of worse attentional abilities in cities suggests that urbanites might have, for example, less effective economic, creative, and social interactions. These conclusions fit with the fact that attentional capabilities do vary contextually between and within individuals: for example, targeted training can improve sustained attention Lutz et al. [2009], Luo and Zhang [2020], there are cultural differences in attentional engagement Kardan et al. [2017], Köster et al.

[2018], Alotaibi et al. [2017], attention can change across the human lifespan Fortenbaugh et al. [2015], and poorer attention is often associated with mental health disorders Clark and Goodwin [2004], Swaab-Barneveld et al. [2000], Prouteau et al. [2004] and low socioeconomic status Razza et al. [2010], Hoyer et al. [2021]. Along these lines, research with small samples has suggested that urbanization decreases attentional engagement Linnell et al. [2013], Caparos et al. [2013], Linnell et al. [2014]. In particular, these studies suggest that when compared to individuals from rural environments, urban inhabitants have higher levels of tonic alertness that favor more diffuse attention across varied stimuli and subsequent poorer attentional engagement on task-specific stimuli. However, larger cities have also been associated with greater economic productivity Bettencourt [2013], more social interactions Schläpfer et al. [2014b], and greater creativity Bettencourt et al. [2007b] which is not consistent with drastically worse attention in cities.

Two reasons for these discrepancies may be that (1) these studies focus primarily on wholesale differences between rural and urban and (2) they apply coarse characterizations of attention despite the fact that attention is multifaceted Rosenberg et al. [2018]. Towards the first point, urban/rural differences do not probe whether there are systematic variations in attention between different urban environments. Such differences in human behavior have recently been demonstrated for mental depression Stier et al. [2021], implicit racial biases Stier et al. [2023], and navigational abilities Coutrot et al. [2022] across city characteristics such as size, demographics, and geographical layout. These studies generally yield a greater mechanistic understanding of how cities engender psychological adaptations when compared to studies contrasting urban and rural environments. Such across-city studies are amenable to modern mathematical urban science models Bettencourt [2021b, 2013] which provide a self-consistent framework for evaluating and proposing mechanisms governing urban life, from street network layouts and human mobility to socio-economic interactions.

Towards the second point, there are multiple types of attention across different domains,

including alertness and vigilance, directing attention towards stimuli, selecting stimuli to attend to, and sustaining attention Rosenberg et al. [2018], Scholl [2009]. This is important because one can have better performance in some domains, but worse performance in others Green and Bavelier [2012], Huang et al. [2012]. While all domains of attention may be relevant to city life, selective attention is the domain most commonly invoked, often implicitly, in existing studies. This literature has not, in general, directly measured attention with cognitive tasks. Consequently, we aim here to directly investigate selective attention beyond urban/rural comparisons and to determine whether it systematically varies between cities with different characteristics.

Results

Greater Selective Attention Performance in Larger Cities

In order to assess systematic differences in selective attention we follow previous studies that focus on population size, i.e., scale, as a primary driver of urban characteristics Bettencourt [2021b, 2013], Bettencourt et al. [2007a], Stier et al. [2021, 2022a, 2023]. The logic of these studies is based on the mathematical models of Urban Scaling Theory Bettencourt [2021b, 2013]. These models describe cities as arising from balancing costs (e.g., transportation, housing, utilities) with the benefits that arise from socioeconomic interactions which occur as people move through cities. Importantly, the models have been extensively validated with empirical data across human history and different cultures Bettencourt [2013, 2021b], Lobo et al. [2020], Ortman et al. [2014b, 2020]. Along with stochastic growth Bettencourt [2020], they describe how the built environment and social functions emerge concomitantly in cities.

The standard version of these models produces scaling relationships in which many urban characteristics change systematically with city size in a scale-invariant way. Importantly, these scaling relationships are analytically derived Bettencourt [2013, 2021b] and have been

extensively validated empirically, i.e. been statistically confirmed to match urban statisticsBettencourt et al. [2007a], Bettencourt [2013], Stier et al. [2021, 2023]. Two examples of these scaling relationships are, per-capita socioeconomic interactions (k) and economic outputs (y) follow scaling relationships of the form $k, y \sim N^{\delta}$, where N is city population size, and $\delta = \frac{1}{6}$ is the scaling exponent derived via Urban Scaling Theory Bettencourt [2013, 2021b]. Similarly, per-capita land area, a, follows a scaling relationship of $a \sim N^{\alpha-1}$, where $\alpha = \frac{2}{3}$. These relationships predict that, on average, larger cities – those with greater population – should have disproportionately more social interactions and economic outputs (e.g., gross domestic product) than smaller cities. In contrast, larger cities should have disproportionately less land area per capita than smaller cities, and consequently, have higher population densities. For this reason scale, i.e., population size N, can be viewed as the first-order driver of many urban characteristics, including population density, interaction quantities, and economic outputs.

With this in mind, we used the core based statistical area (CBSA) definition of cities provided by the United States Office of Budget and Management. CBSAs are functional definitions of cities that capture the spatiotemporally extended social networks of cities beyond “downtowns". As a result these definitions include, suburbs and outlying areas that capture where people live, socialize, and work Stier et al. [2022d].

In order to determine whether attentional capabilities show systematic variation with city size, we measured attention via a dot-tracking task in 3,082 participants (see Methods, Figure 5.1). Briefly, participants were shown ten black dots on their smartphone screen and cued to track two to five of them. After all ten dots moved around the screen for 5 seconds, participants were cued with one of the ten dots and asked to indicate if it was one of the initial targets. Hit rate was the primary outcome measure and chance performance ranged from 0.2 to 0.5 depending on the trial type. This task probes selective attention Sears and Pylyshyn [2000], Scholl [2009], Yoo et al. [2022] which is particularly salient for urban life:

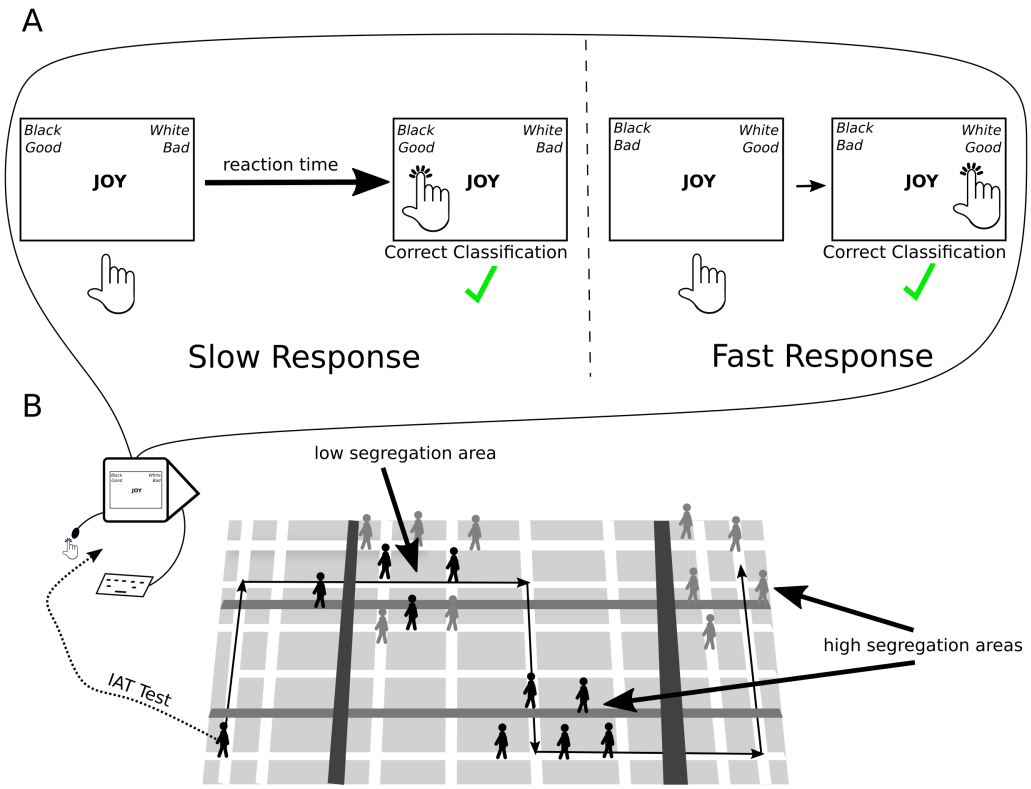


Figure 5.1: A. The dot-tracking task measures selective attention. Participants are cued to track a number of targets. After all targets move around the screen for 5 seconds, they are probed with one dot and asked to identify whether or not it was one of the targets. B. A map of participants' city location. Hawaii, Alaska, and Puerto Rico are excluded for visualization purposes.

consider the attentional demands on a parent with their child and dog trying to safely cross a street while navigating bike lanes and avoiding cars.

First, in order to provide empirical validation of this dataset, we reproduced a recent result that demonstrated scaling of depression rates with city population Stier et al. [2021]. Twenty depression-related questions were included in the demographic questionnaires conducted during the study, which were coded from 1-4, with higher scores indicating a greater likelihood of depressive symptoms. We calculated depression rates as the number of individuals with scores greater than a given threshold. Since there were few repeated samples within all but the largest cities (samples per city: for depression, mean=3.98, max=88, standard deviation=7.66; for attention mean=5.78, max=148, standard deviation=13.10), we used Nadarya-Watson kernel regression Nadaraya [1964], Watson [1964] to efficiently bin observations from cities of similar population sizes in an assumption-free manner (see Methods). We found evidence, across all threshold choices, for lower depression rates in larger cities (Supplementary Figure 7.1 & Supplementary Table 7.1), and scaling exponents consistent with previous evidence Stier et al. [2021]. This result suggests that the dataset and the kernel regression method for combining observations across cities of similar sizes are suitable for investigations of scaling with city population size.

Returning to attention, approximately 7% of participants performed at or near ceiling and a similar proportion performed at or near floor on the dot tracking task. Consequently, standard measures such as hit rate and d' (a standardized measure of both false alarm and hit rate) are not, on their own, particularly useful for understanding performance at the population level. Instead, we calculated the rate of people in different urban environments who were performed at or near ceiling (floor) operationalized as hit rate greater than (less than) a threshold. We additionally examined the rate of individuals with high and low false alarm rates.

We found that larger cities tend to have better attentional performance (Figure 5.2)

and that this result was robust to the exclusion of individuals who indicated they were not paying attention (Supplementary Figure 7.2). In fact, the rate of individuals who indicated that they were distracted decreased in larger cities (Supplementary Figure 7.3). Though this decrease was less dramatic than the observed increase in attention, this suggests that individuals in larger cities may be better at engaging in the dot-tracking task in the first place, and that these results may extend to other domains of attention beyond selectivity. Additional investigations with a variety of attentional tasks will be necessary to confirm which other domains of attention scale with city population.

We also found that the rate of particularly poor performers decreases with city size (Supplementary Tables 7.4 & 7.7), though not as drastically as the rate of particularly good performers. This indicates that the selective attention benefits of larger cities apply broadly, and not just to the best performers. Similarly, we observed a significant, though much smaller increase in the rate of individuals with low false alarm rates (Supplementary Table 7.5) and a decrease in the rate of individuals with high false alarm rates (Supplementary Table 7.6). Together, these results demonstrate that selective attention performance is increased in larger cities and suggest that attentional performance may be increased across other domains as well.

Importantly, these results were not sensitive to individual demographics including income, education, race/ethnicity, age, and self reported health (Supplementary Table 7.2), sleep quality or stress (Supplementary Table 7.3), or the inclusion of data from only participants' first time participating in the dot tracking task (Supplementary Figure 7.4).

Beyond the qualitative result of greater attention in larger cities, we found that the scaling exponent associated with this increase in attentional abilities in larger cities ($\beta = 0.163$; 95% Confidence Interval $[0.132, 0.197]$; for hit rate > 0.9) matched the prediction of $\delta = \frac{1}{6}$ for per-capita social interactions from Urban Scaling Theory Bettencourt [2013, 2021b], Schläpfer et al. [2014b] (see Supplementary Tables 7.8 & 7.9 for sensitivity to the inclusion of smaller

cities and different hit rate thresholds). Though this result was unexpected, it suggests

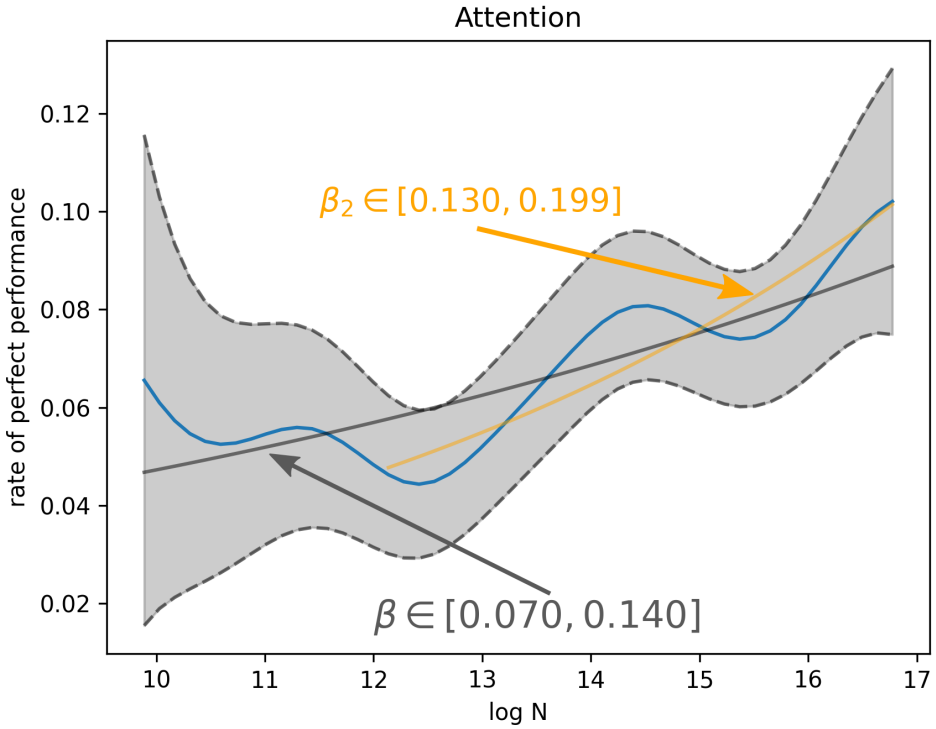


Figure 5.2: Selective attention performance is higher in larger cities. The rate of individuals with near-perfect performance (hit rate $> .9$) increases with city population, N. For the cities where we have a larger amount of data the scaling exponent is consistent with the Urban Scaling Theory prediction of $\beta = \frac{1}{6}$ for increasing social interaction density. The blue line shows the Nadaraya–Watson kernel regression estimate and the envelope with dashed lines shows the 95% confidence interval for the kernel regression.

a fundamental link between selective attention capabilities and social interactions in cities. Further, these results warrant further consideration of whether greater attention results from or drives the scaling of social interactions with city population.

Theoretical Implications of Selective Attention Scaling

To begin to answer this question, we examine Urban Scaling Theory's predictions about social interactions more closely to determine any theoretical constraints on potential relationships between attentional abilities and social interactions.

The starting point for the derivation of the interaction scaling relationship ($k \sim N^{\delta}$) is the assumption that, on average, all individuals have the same probability of interacting with others as they move over a city's infrastructure networks Bettencourt [2013, 2021b]. This probability is further broken down into two parameters, the first is a cross-section a_0 which is the typical radius of interaction: if two individuals come within a distance of a_0 of each other, they will have an interaction. The second parameter is a characteristic length, l, traveled by each city inhabitant during a given time period. Thus, the quantity $\frac{a_0 l}{A_n}$ represents the fraction of the total area of a city's infrastructure network, A_n over which an average individual is available for social interactions in a given time period. In other words, this is the probability of having a social interaction.

To finish the derivation we multiply the above quantity by the number of individuals moving through the infrastructure network, i.e., the city population, N. The average per capita interactions are then:

$$k \sim \frac{a_0 l}{A_n} N \tag{5.1}$$

To see that this results in the typical scaling relationship, we can apply the result that $A_n \sim N^{1-\delta}$ Bettencourt [2013]. This leads to $k \sim \frac{a_0 l}{N^{1-\delta}} N \sim a_0 l N^{\delta}$.

With this derivation in hand, we can now ask whether it is theoretically possible that greater attention (Figure 5.2) could drive the disproportionately greater quantities of interactions observed in larger cities Schläpfer et al. [2014b]. To answer this question, we examine each part of Equation 5.1 in turn.

First, A_n, the area of cities' infrastructure networks is a property of the physical environment and is unlikely to be influenced by the attentional capabilities of urban inhabitants. Similarly, the total population of a city is unlikely to be primarily driven by individual selective attentional capabilities. In addition, Urban Scaling Theory starts with the assumptions that a_0 and l are the same, on average, across all individuals of cities of different sizes. Thus, in the Urban Scaling Theory framework, the observation of disproportionately greater

numbers of interactions in larger cities is purely a built environment effect resulting from realized economies of scale in infrastructure use (e.g., more people travel over the same size road in larger cities).

As a result, if increasing selective attentional abilities were to impact the quantities of social interactions in cities, that effect would have to occur through the quantity $a_0 l$. For the moment we can set aside the psychological validity of such a connection between selective attention and interaction probabilities or characteristic travel distances. If such a connection existed, it would result, generally, in the quantity $a_0 l$ scaling with city population with some scaling exponent different from zero. In particular, a positive relationship between attention and $a_0 l$, i.e., greater selective attentional abilities associated with a higher probability of social interactions, will yield a positive scaling exponent ($a_0 l \sim N^{\xi};\ \xi > 0$), while a negative relationship between attention and $a_0 l$ will yield a negative scaling exponent ($\xi < 0$). This would lead to an interaction scaling relationship of $k \sim a_0 l N^{\delta} \sim N^{\xi} N^{\delta} \sim N^{\delta+\xi}$. For example, in the simple case where $a_0 l$ is directly proportional to selective attention, the result is an interaction scaling relationship of $k \sim a_0 l N^{\delta} \sim N^{\delta} N^{\delta} \sim N^{2\delta}$. This does not match the relationship $k \sim N^{\delta}$ that is observed in empirical data Schläpfer et al. [2014b]. Therefore, from the perspective of Urban Scaling Theory, it is unlikely that selective attention is playing a mechanistic role in driving the scaling of quantities of social interactions with city population.

What role, then, could selective attention play in urban social interactions? One might suggest that the increased attentional demands of larger cities simply act like training and that attentional abilities respond to the increased social stimuli. While this may indeed be the case, the question of how enhanced attention modifies urban behavior remains: what do urban inhabitants get from increased selective attentional abilities and what does it cost them to acquire? The theoretical considerations above suggest that the benefits (or detriments) of increased selective attention cannot simply be more (or fewer) interactions. Instead, we

suggest that selective attentional capabilities may drive social interaction quality in cities.

Though there is little research directly linking selective attention (or other domains of attention) to social interaction quality, the literature contains some hints. One study using functional magnetic resonance imaging (fMRI) in humans has shown that the quality of social interactions is preferentially mapped in the medial prefrontal cortex Peer et al. [2021], an area of the brain that has been associated with attentional broadly Kahn et al. [2012], Rossi et al. [2009], Passetti et al. [2002]. At the same time, there are a number of studies associating greater theory of mind (ToM) abilities (otherwise known as mentalizing) with social network quality and size through attentional processes. For example, greater ToM has been associated with larger numbers of close contacts contacts Schmälzle et al. [2017], Stiller and Dunbar [2007]. Interestingly, worse ToM task performance has been separately associated with attentional performance deficits, particularly in executive control aspects of attention Mary et al. [2016], Lin et al. [2010], Leslie et al. [2004], Tatar and Cansız [2022]. Beyond these studies, the literature has focused on particularly poor attentional performance associated with attention-deficit/hyperactivity disorder (ADHD). These studies demonstrate that ADHD and related deficits in attention are associated broadly with poorer quality social interactions Hoza et al. [2005], Greene et al. [2001], Normand et al. [2013], Kim et al. [2015]. This evidence suggests that greater selective attention performance in larger cities may result in individuals in larger cities having, on average, higher quality social interactions.

However, the drastic attentional deficits associated with ADHD may not be particularly relevant to the smaller shifts in selective attention between cities that we report here (Figure 5.2). In particular, while the literature paints a picture of poorer attention leading to poorer interaction quality, some individuals with greater attentional capacity might choose to spend this increased attentional budget on a wider number of weak ties. Such weak ties have long been considered to promote a number of positive qualities including well-being, job mobility, and collective efficacy Granovetter [1973], Rajkumar et al. [2022], Kavanaugh et al. [2003].

Towards, this point, recent research on online social networks has suggested that individuals devote the most time to interactions with weak and very strong ties Weng et al. [2018]. These types of weak relationships may be particularly relevant in large cities where divisions of labor encourage reliance on a loose network of interactions, especially for services (e.g., in a large city one's plumber, electrician, and roofer are distinct individuals). Indeed, recent evidence from Chilean phone networks Samaniego et al. [2020] and United States social media networks (Supplementary Figure 7.5) supports this idea. Those analyses reveal that social network clustering (the fraction of contacts who are themselves connected) decreases as cities get larger. This suggests that up to a point (see Samaniego et al. [2020] and Supplementary Figure 7.5), people in larger cities have a greater quantity of weak ties and less close-knit social networks. Thus, greater attentional abilities might result in a decrease in the quality of an individual's interactions, on average, while still being associated with beneficial outcomes.

Discussion

To summarize, we found evidence that selective attention scales with city population size such that individuals in larger cities have disproportionately better selective attention. While we found this effect across different measures of attention including self-reported task engagement, false alarm rates, and hit rates, the strongest effect was for hit rates (which emphasize accurate selective attention). Further, the scaling exponent we observed for selective attention and city population matched that predicted by Urban Scaling Theory for social interactions. This suggests a fundamental link between attention and social interactions in cities.

Based on these results, a review of the Urban Scaling model revealed that, under this framework, it is unlikely that better selective attention in larger cities is driving greater numbers of per capita interactions. Instead, we suggest that greater selective attention performance drives systematic differences in social interaction quality between cities. Our

empirical data precluded confirming this hypothesis or clarifying how increased selective attention may benefit (or disadvantage) inhabitants of larger cities. While existing psychological and neuroimaging evidence suggests that greater attentional performance may result in higher-quality interactions, evidence from sociology and urban science suggests that larger cities may foster more sparse social networks with lower-quality interactions.

This mismatch between the ecological, population-level perspective and the individual perspective warrants further investigation. The population-level observations may suffer from an ecological fallacy Loney and Nagelkerke [2014], in which the direction of the observed population-level effect is not indicative of the underlying mechanism, but rather an additional confounding variable (e.g., in a Simpson's Paradox). However, existing evidence might also be inadequate to draw conclusions regarding the relationship between selective attention and interaction quality in cities. Thus, the discrepancy between these two perspectives may be resolved by the simultaneous collection of explicit attentional measures and measures of social network structure and tie quality.

The results here are situated in the context of Urban Scaling Theory Bettencourt [2013, 2021b]. The majority of similar theoretical and empirical work has focused on what can be learned by examining how quantities of interactions systematically vary with various city characteristics. However, here we suggest that in addition to quantity, the quality of social and economic interactions may be particularly important to consider when attempting to construct more detailed models of urban life. This is an open frontier in urban science, and more work is needed to elucidate what mechanisms drive urban interaction quality. In addition, the fundamental role of attention in social interactions has seen little explicit investigation in the psychological literature (beyond how extreme deficits associated with mental health disorders impact sociality). Thus, further investigations of interaction quality in cities have the potential to advance basic psychological understanding of human attention and sociality.

We conclude by noting that cities provide an excellent balance between constraints on and diversity of human behavior. For one, cities are restrictive: inhabitants must traverse the city within the infrastructure networks and are forced to interact with others in ways that they might not choose to in a rural area. At the same time, these interactions engender a proliferation of diversity in culture and opportunity. The constraints imposed by cities allow us (researchers) to average over some of the stochasticity in human behavior when considering large urban populations, and the diversity of urban social and economic functions provides enough variance to yield rich insights into human behavior. Thus, cities are well suited to ecologically valid investigations of human behavior, and investigations of the type we present here show great promise to help us better understand cities specifically, and human behavior generally.

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.

This work is partially supported by NSF-2106013.

Materials and Methods

Attention and Depression Data

Attention data were collected as part of a partnership between the University of California, San Francisco, and Samsung Research America that aimed to examine the impact of daily psychological stress on various indicators of health. The study ran from March 2018 until December 2021, with cognitive task assessments included from October 2019 onwards. Participants were enrolled in a 21-day research study via the MyBpLab app on Samsung devices. Overall, 164,759 people worldwide participated in the study. Participants provided demographic information upon enrollment (e.g., zip code or country, age, sex, etc.) and some participants also completed additional surveys, for example, a depression screening questionnaire. All participants who completed the dot-tracking attention task did so at night (6-10 pm) on days 2, 7, 11, 16, & 20. A total of 14,489 participants completed at least one dot-tracking session, though only 8,507 of those participants had zip code information available. In addition, immediately after each task session, participants were asked about the degree to which they felt distracted during the task. Only 3,082 participants reported not being distracted at all during the task. Of the full study sample (164,759) only 1,806 participants completed the depression questionnaire and provided zip code information.

Population Data

All census data is publicly available and was downloaded from `data.census.gov`. Five-year population estimates for U.S. cities defined as combined statistical areas (CBSAs) were downloaded from table *B01003*. In order to map between zip codes and CBSAs, crosswalk files for 2021 were downloaded from the United States Department of Housing and Urban Development's Office of Policy Development and Research from `https://www.huduser.gov/portal/datasets/usps_crosswalk.html`.

Nadarya-Watson Kernel Regression

Kernel regression was used to combine observations from cities of similar sizes in a non-parametric manner. The dependent variable used in the kernel regression was a binary variable indicating whether the measure of interest (hit rate or depression score) was greater than (or less than) a given threshold. The independent variable was the natural logarithm of the city population for each participant. Thus, the kernel regression estimates returned the mean rate of the measure being greater than (or less than) a given threshold). The R library *np* was used to implement the kernel regression. This library provides functions for calculating means as well as standard errors, the latter of which was used to construct 95% confidence intervals.

Attention and Depression Scaling Analysis

In order to determine scaling exponents for Attention and Depression we employed ordinary least squares regression (OLS) using the outputs for the kernel regression. In order to obtain proper parameter estimate confidence intervals, we follow the following procedure. First, for each output bin in the kernel regression, we sampled a new value for the independent variable from a normal distribution with a mean equal to the original kernel regression mean and a variance derived from the kernel regression 95% confidence interval as $\sigma^2 = (y_{max} - y_{mean})/PPF(0.975)$, where y_{max} is the 95% confidence interval upper bound, y_{mean} is the kernel regression mean estimate and PPF is the standard normal distribution percentage point function. The parameters were then re-estimated from these samples with OLS. This procedure was carried out 1,000 times and parameter estimate confidence bounds were obtained as the 2.5% and 97.5% percentiles of these 1,000 estimates.

Clustering Data and Scaling Analysis

We replicated a previous result from Chilean cell phone networks Samaniego et al. [2020] that clustering (defined as the fraction of contacts who are themselves contacts) decreases with city population size. In particular, those researchers found that while clustering decreased overall, there seemed to be two regimes. First, for small cities, clustering decreased with city population. Second, for larger cities clustering was relatively constant with city population.

We replicated this in a large dataset derived from online social network connections from Meta that had been previously published and released publicly Chetty et al. [2022a,b]. These data were downloaded from *https://socialcapital.org/* and contained clustering coefficient estimates for each zip code in the United States. In order to aggregate from zip codes to cities, we took a population-weighted average of all of the zip codes contained in each core-based statistical area.

We then fit scaling relationships for clustering with OLS to (1) the entire dataset, (2) small cities, and (3) large cities. In order to find the cutoff point for small and large cities, we first assumed that there would be two different scaling regimes. Next, we picked a particular threshold to separate the two regimes and performed separate OLS regressions for each regime. Over many choices of threshold, we calculated to total prediction error across both OLS models (above and below the threshold) and then chose the threshold that resulted in the lowest prediction error. This resulted in a threshold of approximately 124,500 separating large and small cities.

CHAPTER 6

DISCUSSION

The goal of the research presented here was to work towards a new framework for broadly understanding human behavior and to do this in an ecologically relevant way. We've seen how cities can be an excellent tool for psychological study. In particular, the fact that they constrain people and force interactions creates regularities that allow us to study urban behavior scientifically (these constraints give cities some qualities in common with psychological laboratories). On the other hand, cities also provide enough diversity of human behavior that they are interesting: they provide a broad assay of human behavior. The city, as a psychological laboratory, is far removed from traditional research settings in which we isolate very specific behaviors. This brings benefits of ecological validity and the ability to study human behavior and the consequences of group dynamics at scale. However, the trade-offs are a potential lack of precision and a definite inability to perform the same kind of causal, controlled experiments that we are used to in traditional psychology laboratories.

Nonetheless, the results of applying this new lens for observing and understanding human behavior have been surprising at times. In Chapters 2 and 5 we saw that cities foster lower depression rates and better selective attention, in contrast to previous assumptions and intuitions based on projections from smaller-scale studies. On the other hand, unsurprisingly, Chapters 3 and 4 demonstrated ways in which segregation and a lack of diversity is problematic. Though the negative impacts of segregation and low diversity have been well established previously, the point here was to precisely predict its impacts, to provide a mechanistic account of how segregation may drive worse outcomes for people in different cities, and to situate these effects in a wider theory.

This last point is a particularly important one. We can, and do, employ many different models to study human behavior across disciplines, from agent-based (e.g., Schelling [1971]), to more simple analytic models (Bettencourt [2013, 2020, 2021b]), to psychological process

models (Dalege et al. [2017]). However, it generally has not been clear how to integrate these results. One possibility is that this is not possible: that the diversity and complexity of human behavior is so great that there is no real synthesis or general theoretical work to be done.

I disagree. The research presented here is a start towards integrating investigations of human behavior with mathematical theory, particularly at scale. Already, we have seen that integrating the bottom-up psychological perspective with the top-down urban science approach has brought a greater understanding of human psychology and urban dynamics. These successes are expected to continue as Urban Psychology continues to reveal what the psychology and urban science literatures can teach one another more broadly.

The specific approach I took was to start at the scale of entire cities and the models of Urban Scaling Theory Bettencourt [2021a]. These models average out almost all of the variety in human behavior, though they leave just enough detail to tell the broad strokes story of cities. The way forward, which has begun with the research in Chapters 3,4, and 5, is to engage in a slow process of de-averaging. What I mean by this is precisely to start with simple models that, for example in the case of the standard Urban Scaling Theory, treat all individuals in the city as identical, everyone is average. The hard work, then, is to add more detail, little by little, de-averaging as necessary to explain increasingly complex behavior with as simple a model as possible.

While this may not always work, and analytically tractable models may not always be possible, there is great promise in this approach. The best example of this here is the addition of group membership and heterogeneous mixing to the standard Urban Scaling Theory model (Chapters 3 and 4). We saw that when studying implicit biases, these simple considerations of more complex interactions that de-average only to the point of a small number of groups and simple interaction segregation, were particularly successful at explaining empirical variance in implicit racial biases.

These investigations succeeded by (1) identifying important features of the environment to which people are adapting and (2) proposing mechanisms of adaptation based on existing psychological and biological findings. One path forward is to continue in this vein (empirically and theoretically) by further enumerating the consequences of mechanisms that mediate the ecological effects of cities: social network size, exposure processes, and learning processes. While this approach has great promise, it will, at some time or another, come to an impasse and a return to the laboratory will be necessary. This is already close to becoming necessary to start to better understand the patterns and mechanisms relating selective attention and social network quality in cities.

In Chapter 5, we demonstrated that larger cities are associated with better performance on selective attention tasks and suggested that this might be related to different social interaction qualities between larger and smaller cities. Even once it is known whether these differences in attention are related to higher-quality or lower-quality interactions, the circular feedback loops that exist in cities preclude making direct hypotheses about interventions. For example, would attentional training impact social interaction quality differently in a small vs. large city? Such questions could be answered directly by recruiting participants from different cities and comparing results. However, this question, and other similar ones may also benefit from laboratory-based experiments. For example, participants with various levels of attentional performance could be recruited to engage in collaborative online video games that share similarities to city life. This would allow direct experimental manipulation of the "attentiveness" of a virtual city and provide a basis for direct causal conclusions. It is likely that such experimental work will become integral to Urban Psychology as it discovers new psychological and urban phenomena.

I think it fitting to conclude with a quote from Jane Jacob's "The Death and Life of Great American Cities". In many ways, her work has been an inspiration and source of wisdom for modern urban science. I also consider her work an inspiration for Urban Psychology which

promises new understanding of human behavior. Although the research presented here is just the beginning, it gives us a concrete way to "start, if only in a small way, adventuring in the real world, ourselves. The way to get at what goes on in the seemingly mysterious and perverse behavior of cities is ... to look closely, and with as little previous expectation as is possible, at the most ordinary scenes and events, and attempt to see what they mean and whether any threads of principle emerge among them" Jacobs [2016].

CHAPTER 7

APPENDIX

7.1 Appendix A

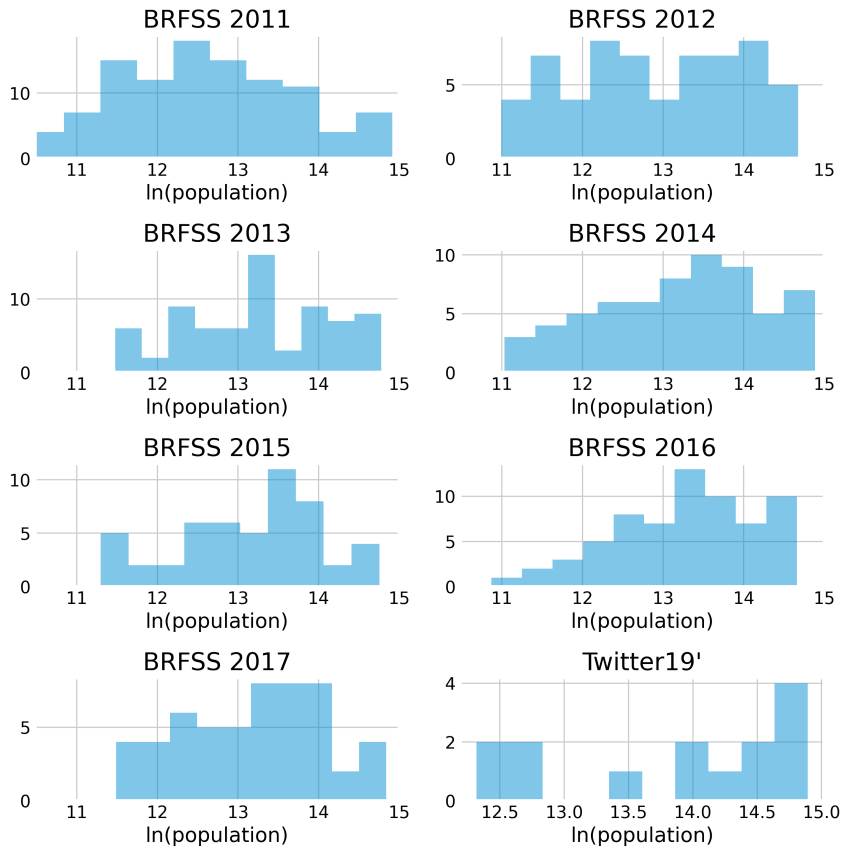


Figure 7.1: Histograms of the detected change points for all window sizes in BRFSS data and Twitter19'. We used a covariate discriminant method (see Methods) to non-parametrically detect changes in the joint distribution of depression rates and population, under the assumption that BRFSS report methods might induce an artificial change. For the Twitter19' dataset, the detection of change points primarily at the edges of the population range is indicative of finite edge effects rather than a true change in the joint distribution of depression rates and city size. Compare to BRFSS data where change points are detected in the middle of the population range.

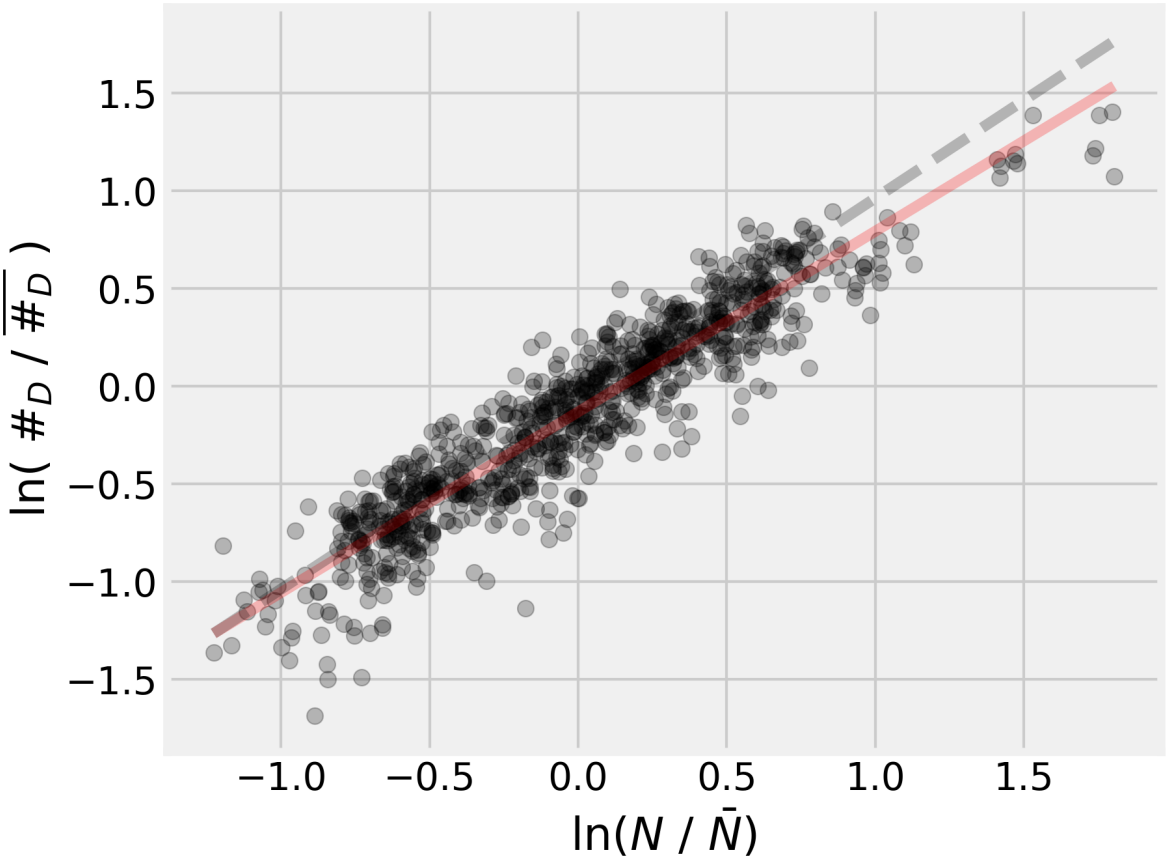


Figure 7.2: Pooling BRFSS data across years for all cities results in a scaling exponent of $\beta = 0.926$ (95% CI = [0.903, 0.950]), consistent with lower depression rates in larger cities.

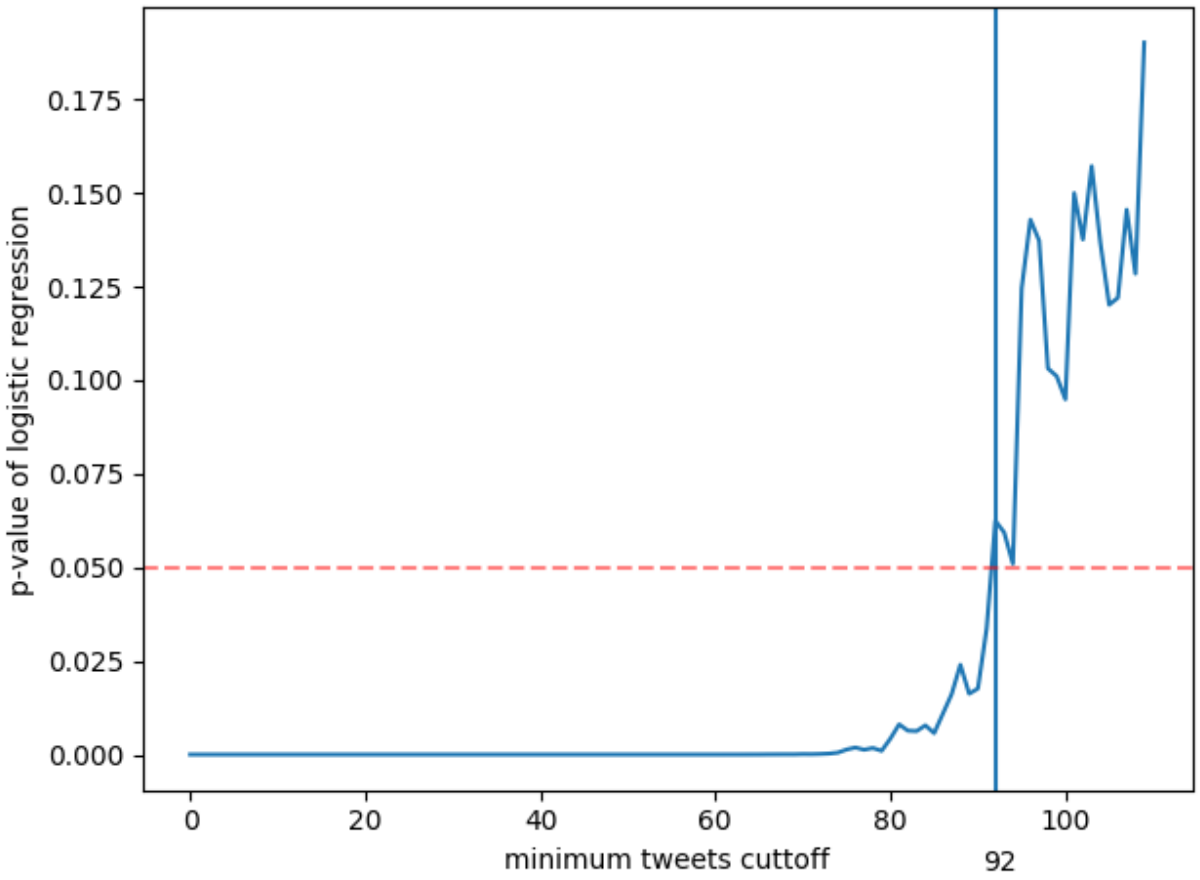


Figure 7.3: Users with lower numbers of tweets are more likely to have depressive sentiment in their tweets. When using an exclusion criteria of less that 92 tweets a logistic regression model significantly distinguishes individuals with depressive sentiment from individuals without depressive sentiment.

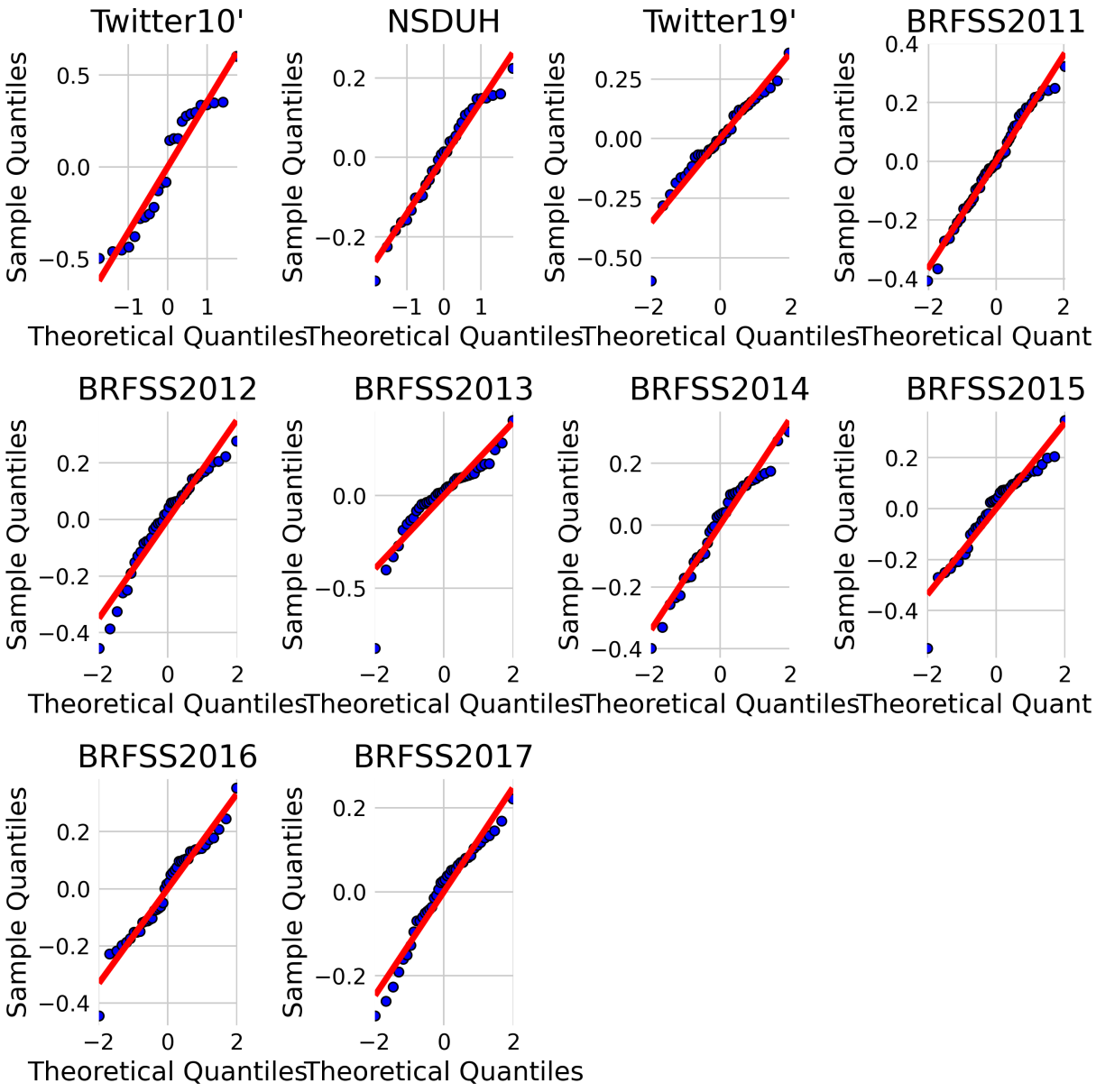


Figure 7.4: QQ plots of the residuals of the OLS model. No significant deviations are observed indicating that the residuals are approximately normally distributed and the linear model is appropriate.

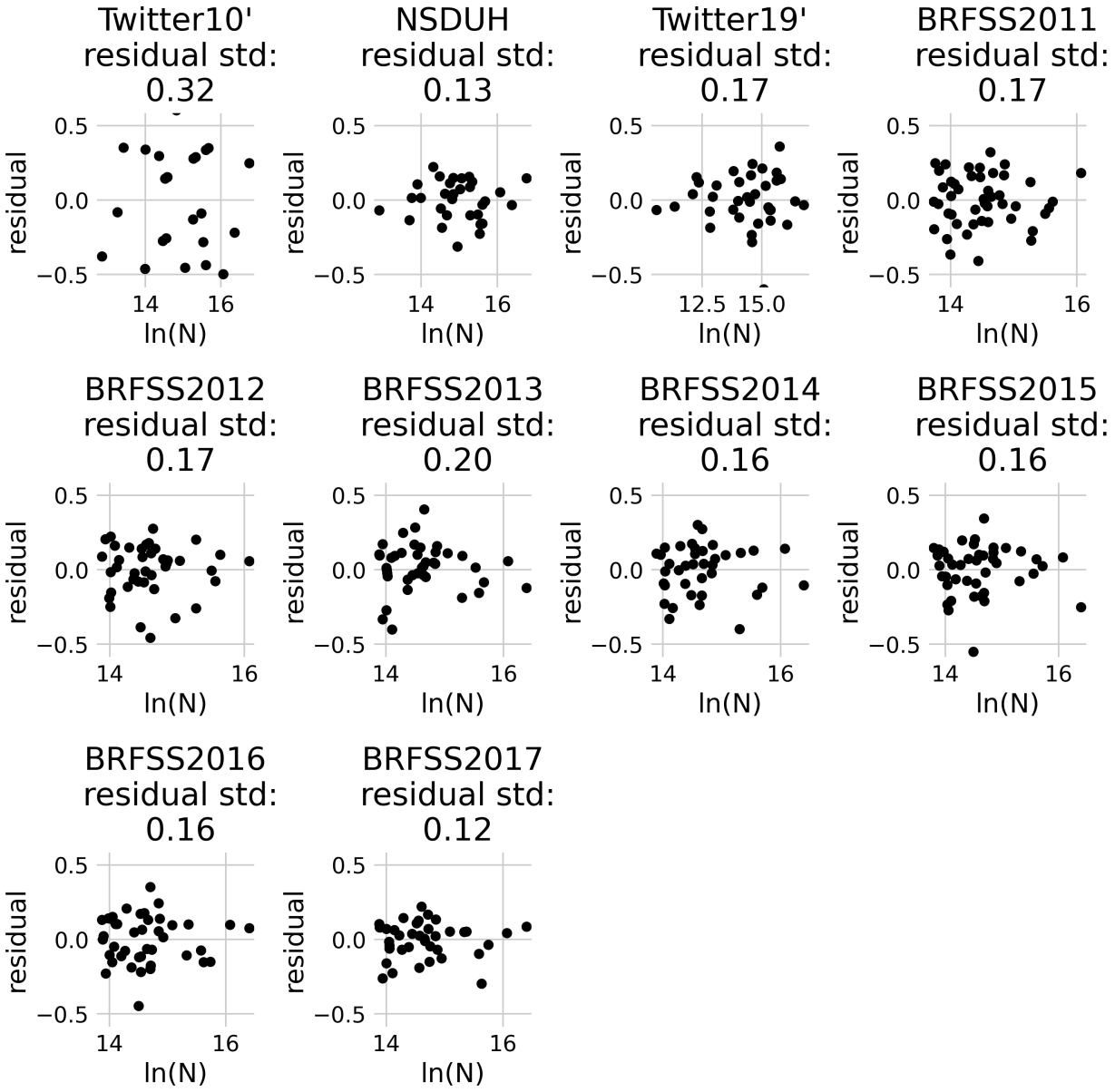


Figure 7.5: Residuals from OLS models are not correlated with city size. In all datasets, residuals are not correlated with city size (Spearman-r minimum p-value = 0.44). Thus no corrections to estimates of β are required.

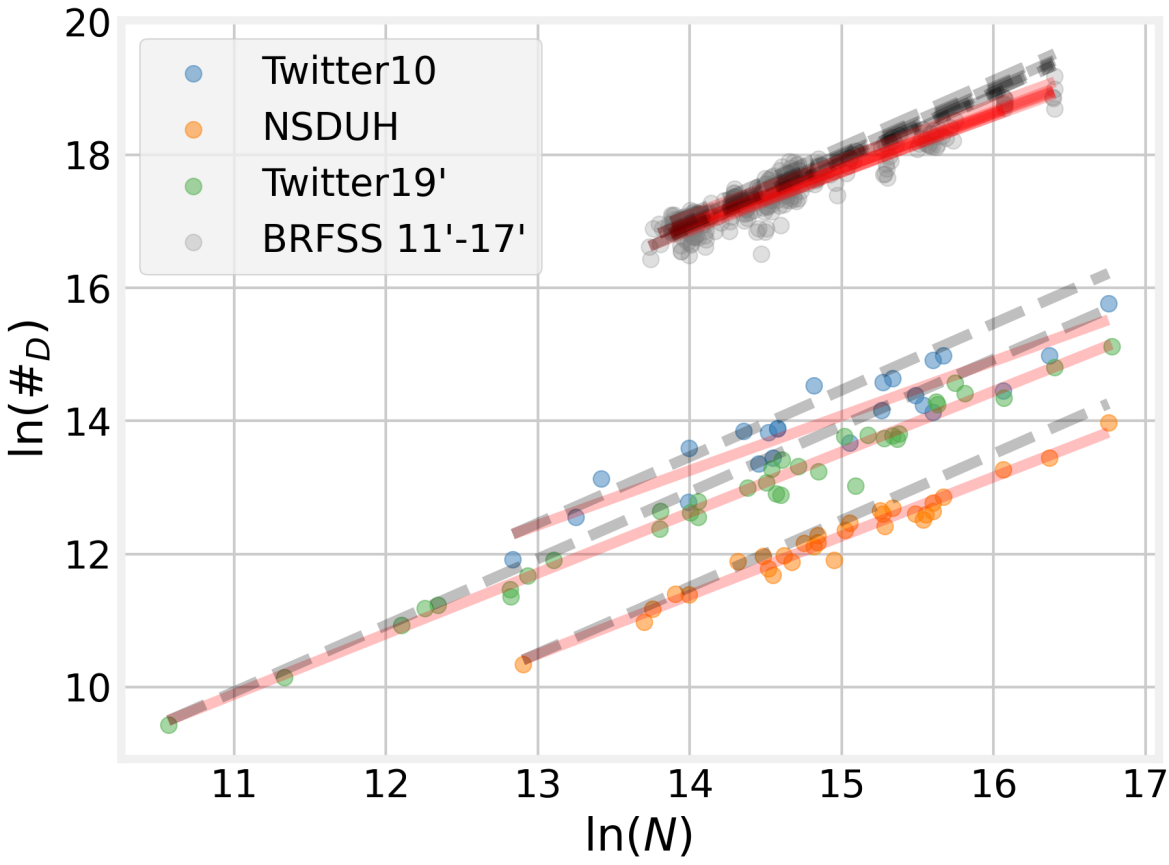


Figure 7.6: OLS fit to each dataset. Sublinear scaling is observed across all datasets.

Table 7.1: MSAs included in the analysis in the main text (Fig. 2.). Included MSAs are marked with an X. The Twitter datasets are abbreviated to Tw.

MSA	Tw10'	BFRSS	NSDUH	Tw19'	2017 Pop
Albuquerque, NM	-	-	X	-	912897
Ann Arbor, MI	-	-	-	X	369208
Atlanta-Sandy Springs-Alpharetta, GA	X	X	X	-	5874249
Augusta-Richmond County, GA-SC	X	-	-	-	600006
Austin-Round Rock-Georgetown, TX	-	X	-	-	2115230
Baltimore-Columbia-Towson, MD	X	X	X	-	2798587
Birmingham-Hoover, AL	-	X	-	-	1085750
Boston-Cambridge-Newton, MA-NH	X	-	X	-	4844597
Bridgeport-Stamford-Norwalk, CT	-	X	-	-	943457
Buffalo-Cheektowaga, NY	-	X	-	-	1129660
Charlotte-Concord-Gastonia, NC-SC	-	X	-	-	2549741
Chicago-Naperville-Elgin, IL-IN-WI	X	X	X	X	9520784
Cincinnati, OH-KY-IN	X	X	-	X	2202597
Cleveland-Elyria, OH	X	X	X	X	2058549

Continued on next page

Table 7.1: MSAs included in the analysis in the main text (Fig. 2.). Included MSAs are marked with an X. The Twitter datasets are abbreviated to Tw. (Continued)

MSA	Tw10'	BFRSS	NSDUH	Tw19'	2017 Pop
Columbus, OH	-	X	-	-	2082475
Dallas-Fort Worth-Arlington, TX	X	-	X	X	7340943
Denver-Aurora-Lakewood, CO	-	X	X	-	2892979
Detroit-Warren-Dearborn, MI	X	-	X	X	4321704
Fresno, CA	-	-	-	X	986542
Grand Rapids-Kentwood, MI	-	X	-	-	1063926
Gulfport-Biloxi, MS	-	-	-	X	412946
Hartford-East Hartford-Middletown, CT	-	X	-	-	1206719
Houma-Thibodaux, LA	-	-	-	X	209893
Houston-The Woodlands-Sugar Land, TX	X	X	X	X	6905695
Indianapolis-Carmel-Anderson, IN	X	X	-	-	2026723
Jacksonville, FL	-	X	-	-	1504841
Kansas City, MO-KS	X	X	X	X	2127259
Lafayette, LA	-	-	-	X	490107
Las Vegas-Henderson-Paradise, NV	-	X	X	X	2183310
Lock Haven, PA	-	-	-	X	38837

Continued on next page

Table 7.1: MSAs included in the analysis in the main text (Fig. 2.). Included MSAs are marked with an X. The Twitter datasets are abbreviated to Tw. (Continued)

MSA	Tw10'	BFRSS	NSDUH	Tw19'	2017 Pop
Los Angeles-Long Beach-Anaheim, CA	X	X	X	X	13298709
Louisville/Jefferson County, KY-IN	-	X	-	-	1260391
Macon-Bibb County, GA	-	-	-	X	229081
Manchester-Nashua, NH	-	-	X	-	413157
Memphis, TN-MS-AR	-	X	-	-	1339290
Miami-Fort Lauderdale-Pompano Beach, FL	X	X	X	X	6149687
Milwaukee-Waukesha, WI	-	X	-	-	1575151
Minneapolis-St. Paul-Bloomington, MN-WI	-	X	X	X	3577765
Montgomery, AL	X	-	-	-	374042
Nashville-Davidson–Murfreesboro–Franklin, TN	-	X	X	-	1875736
New Orleans-Metairie, LA	X	X	X	X	1270465
New York-Newark-Jersey City, NY-NJ-PA	X	-	X	X	19325698
Oklahoma City, OK	-	X	-	-	1383249
Opelousas, LA	-	-	-	X	83447

Continued on next page

Table 7.1: MSAs included in the analysis in the main text (Fig. 2.). Included MSAs are marked with an X. The Twitter datasets are abbreviated to Tw. (Continued)

MSA	Tw10'	BFRSS	NSDUH	Tw19'	2017 Pop
Orlando-Kissimmee-Sanford, FL	X	X	-	-	2512917
Philadelphia-Camden-Wilmington	X	-	X	X	6078451
Phoenix-Mesa-Chandler, AZ	-	X	X	X	4761694
Pittsburgh, PA	-	X	X	-	2330283
Portland-Vancouver-Hillsboro, OR-WA	-	X	X	X	2456462
Poughkeepsie-Newburgh-Middletown, NY	X	-	-	-	673253
Providence-Warwick, RI-MA	-	X	-	-	1617057
Raleigh-Cary, NC	-	X	-	-	1334342
Richmond, VA	X	X	-	X	1269478
Riverside-San Bernardino-Ontario, CA	X	X	-	X	4570427
Rochester, NY	-	X	-	-	1071589
Sacramento-Roseville-Folsom, CA	-	X	-	-	2320381
Salt Lake City, UT	-	X	X	X	1205238
San Antonio-New Braunfels, TX	-	X	-	-	2474274
San Diego-Chula Vista-Carlsbad, CA	-	X	X	X	3325468

Continued on next page

Table 7.1: MSAs included in the analysis in the main text (Fig. 2.). Included MSAs are marked with an X. The Twitter datasets are abbreviated to Tw. (Continued)

MSA	Tw10'	BFRSS	NSDUH	Tw19'	2017 Pop
San Francisco-Oakland-Berkeley, CA	-	X	X	X	4710693
San Jose-Sunnyvale-Santa Clara, CA	-	X	-	X	1993582
Seattle-Tacoma-Bellevue, WA	X	-	X	X	3884469
St. Louis, MO-IL	-	X	X	X	2805850
Tampa-St. Petersburg-Clearwater, FL	-	X	X	-	3091225
Trenton-Princeton, NJ	-	-	-	X	368602
Tucson, AZ	-	X	-	-	1027502
Tulsa, OK	-	X	X	X	991610
Virginia Beach-Norfolk-Newport News, VA-NC	X	X	-	X	1761305
Warner Robins, GA	-	-	-	X	180019
Washington-Arlington-Alexandria	-	-	X	-	6213246
Worcester, MA-CT	-	X	-	-	942303

Table 7.2: Estimates of the scaling exponent made with BRFSS data from smaller cities that were below the estimated change point for each year.

Dataset	β	95% CI	R^2	n
BRFSS2011	1.000	[0.960, 1.039]	0.952	128
BRFSS2012	1.001	[0.961, 1.040]	0.954	122
BRFSS2013	1.020	[0.969, 1.070]	0.953	81
BRFSS2014	1.034	[0.991, 1.077]	0.969	74
BRFSS2015	1.044	[0.996, 1.093]	0.966	67
BRFSS2016	0.967	[0.906, 1.028]	0.931	76
BRFSS2017	1.010	[0.962, 1.058]	0.959	76

Table 7.3: Robustness of scaling exponent estimates made with BRFSS data to variation in the city size below which data was excluded.

Dataset	β	95% CI
BRFSS2011	0.88	[0.87, 0.89]
BRFSS2012	0.85	[0.85, 0.87]
BRFSS2013	0.86	[0.85, 0.87]
BRFSS2014	0.83	[0.83, 0.84]
BRFSS2015	0.83	[0.82, 0.84]
BRFSS2016	0.83	[0.82, 0.84]
BRFSS2017	0.83	[0.83, 0.85]

Table 7.4: Scaling exponent estimates for all BFRSS data. No cities below the change point are excluded.

Dataset	β	95% CI	R^2	n
BRFSS2011	0.966	[0.942, 0.991]	0.974	172
BRFSS2012	0.956	[0.931, 0.982]	0.972	161
BRFSS2013	0.951	[0.920, 0.982]	0.968	122
BRFSS2014	0.959	[0.932, 0.987]	0.978	111
BRFSS2015	0.961	[0.932, 0.990]	0.976	109
BRFSS2016	0.941	[0.910, 0.972]	0.968	119
BRFSS2017	0.965	[0.939, 0.991]	0.980	116

Table 7.5: Robustness of scaling exponent estimates to variation in the minimum number of tweets required for inclusion in the Twitter analyses.

Minimum Tweets	β	95% CI	# MSAs
82	0.85	[0.75, 0.96]	31
83	0.85	[0.75, 0.95]	29
84	0.86	[0.75, 0.96]	29
85	0.87	[0.75, 0.98]	28
86	0.86	[0.75, 0.98]	28
87	0.83	[0.69, 0.97]	26
88	0.83	[0.68, 0.98]	25
89	0.80	[0.65, 0.95]	25
90	0.79	[0.63, 0.94]	24
91	0.80	[0.65, 0.96]	24
92	0.82	[0.67, 0.97]	24
93	0.85	[0.70, 0.99]	22
94	0.84	[0.69, 0.98]	22
95	0.83	[0.70, 0.95]	22
96	0.83	[0.70, 0.97]	22
97	0.84	[0.71, 0.98]	22
98	0.86	[0.71, 1.00]	22
99	0.84	[0.70, 0.98]	22
100	0.81	[0.65, 0.97]	22
101	0.86	[0.68, 1.04]	21

Table 7.6: Shapiro-Wilk test of normality on the OLS residuals for each dataset. The residuals from the BRFSS 2013 data fail this normality test due to one outlier city with a negative residual.

Dataset	statistic	p-value	n
Twitter10'	0.917	5.03e-02	24
NSDUH	0.970	5.26e-01	31
Twitter19	0.948	9.26e-02	36
BRFSS2011	0.977	4.79e-01	48
BRFSS2012	0.951	8.96e-02	39
BRFSS2013	0.873	3.47e-04	40
BRFSS2014	0.964	2.54e-01	38
BRFSS2015	0.951	7.38e-02	41
BRFSS2016	0.969	2.95e-01	43
BRFSS2017	0.959	1.55e-01	40

Table 7.7: Result of logistic regression models for each year of BRFSS data.
We conditioned on log-population, the rate of population change from the previous year, income, race, and education. The income variable had 6 levels baselined by missing, followed by levels from less than $15k to greater than $50k. The education variable had 5 levels baselined by not reported followed by levels from no high-school to graduated college. The race variable had 4 levels with a baseline of White followed by: Black, Asian, and other/multi-racial.

	(2017)	**(2016)**	**(2015)**	**(2014)**	**(2013)**	**(2012)**	**(2011)**
logpop	-0.104^{***}	-0.115^{***}	-0.115^{***}	-0.108^{***}	-0.124^{***}	-0.092^{***}	-0.118^{***}
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
inc1	1.062^{***}	1.098^{***}	1.103^{***}	1.125^{***}	1.102^{***}	1.101^{***}	1.111^{***}
	(0.033)	(0.032)	(0.033)	(0.032)	(0.031)	(0.034)	(0.031)

Continued on next page

Table 7.7: Result of logistic regression models for each year of BRFSS data. We conditioned on log-population, the rate of population change from the previous year, income, race, and education. The income variable had 6 levels baselined by missing, followed by levels from less than $15k to greater than $50k. The education variable had 5 levels baselined by not reported followed by levels from no high-school to graduated college. The race variable had 4 levels with a baseline of White followed by: Black, Asian, and other/multi-racial. (Continued)

	(2017)	**(2016)**	**(2015)**	**(2014)**	**(2013)**	**(2012)**	**(2011)**
inc2	0.583^{***}	0.613^{***}	0.679^{***}	0.598^{***}	0.588^{***}	0.601^{***}	0.568^{***}
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.032)	(0.030)
inc3	0.321^{***}	0.288^{***}	0.344^{***}	0.270^{***}	0.282^{***}	0.292^{***}	0.318^{***}
	(0.033)	(0.034)	(0.034)	(0.034)	(0.034)	(0.036)	(0.034)
inc4	0.206^{***}	0.174^{***}	0.188^{***}	0.157^{***}	0.189^{***}	0.144^{***}	0.164^{***}
	(0.031)	(0.031)	(0.032)	(0.032)	(0.032)	(0.035)	(0.033)
inc5	-0.093^{***}	-0.130^{***}	-0.055^{**}	-0.114^{***}	-0.086^{***}	-0.138^{***}	-0.112^{***}
	(0.024)	(0.025)	(0.025)	(0.026)	(0.027)	(0.029)	(0.027)
edu1	0.626^{***}	0.416^{***}	0.786^{***}	0.434^{***}	0.460^{***}	0.691^{***}	0.167

Continued on next page

Table 7.7: Result of logistic regression models for each year of BRFSS data. We conditioned on log-population, the rate of population change from the previous year, income, race, and education. The income variable had 6 levels baselined by missing, followed by levels from less than $15k to greater than $50k. The education variable had 5 levels baselined by not reported followed by levels from no high-school to graduated college. The race variable had 4 levels with a baseline of White followed by: Black, Asian, and other/multi-racial. (Continued)

	(2017)	**(2016)**	**(2015)**	**(2014)**	**(2013)**	**(2012)**	**(2011)**
	(0.159)	(0.151)	(0.172)	(0.105)	(0.129)	(0.168)	(0.136)
edu2	0.541***	0.238	0.686***	0.302***	0.302**	0.587***	0.024
	(0.157)	(0.149)	(0.170)	(0.102)	(0.128)	(0.167)	(0.135)
edu3	0.755***	0.450***	0.923***	0.511***	0.551***	0.779***	0.262*
	(0.157)	(0.149)	(0.170)	(0.102)	(0.127)	(0.167)	(0.134)
edu4	0.534***	0.261*	0.718***	0.320***	0.345***	0.618***	0.123
	(0.157)	(0.149)	(0.170)	(0.102)	(0.127)	(0.167)	(0.135)
rac2	−0.474***	−0.509***	−0.487***	−0.460***	−0.497***	−0.513***	−0.497***
	(0.027)	(0.028)	(0.029)	(0.027)	(0.027)	(0.029)	(0.028)

Continued on next page

Table 7.7: Result of logistic regression models for each year of BRFSS data. We conditioned on log-population, the rate of population change from the previous year, income, race, and education. The income variable had 6 levels baselined by missing, followed by levels from less than $15k to greater than $50k. The education variable had 5 levels baselined by not reported followed by levels from no high-school to graduated college. The race variable had 4 levels with a baseline of White followed by: Black, Asian, and other/multi-racial. (Continued)

	(2017)	**(2016)**	**(2015)**	**(2014)**	**(2013)**	**(2012)**	**(2011)**
rac3	0.116^{*}	0.108	-0.017	0.029	-0.014	-1.186^{***}	-1.084^{***}
	(0.063)	(0.068)	(0.069)	(0.068)	(0.069)	(0.097)	(0.086)
rac4	-0.327^{***}	-0.371^{***}	-0.392^{***}	-0.336^{***}	-0.335^{***}	-0.069^{**}	-0.074^{**}
	(0.029)	(0.030)	(0.032)	(0.031)	(0.031)	(0.031)	(0.030)
pop	-1.194	-1.803^{*}	-1.427	-1.874^{*}	-1.972	-2.987^{**}	0.718
change	(1.094)	(0.944)	(1.017)	(1.087)	(1.274)	(1.302)	(1.281)
Constant	-2.083^{***}	-1.949^{***}	-2.340^{***}	-1.937^{***}	-1.946^{***}	-2.260^{***}	-1.814^{***}
	(0.157)	(0.148)	(0.170)	(0.101)	(0.127)	(0.166)	(0.134)

Table 7.8: Depression rates are not associated with year over year population change. Results from ordinary least squares fits with the rate of population change included.

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
Twitter 2010 Ln Population	-0.2284	0.107	-2.127	0.045	-0.452	-0.005
Twitter 2010 Population Change %	-184.8047	287.438	-0.643	0.527	-782.564	412.955
Twitter 2019 Ln Population	-0.0846	0.022	-3.842	0.001	-0.129	-0.040
Twitter 2019 Population Change %	-3.0574	4.446	-0.688	0.496	-12.102	5.987

Continued on next page

Table 7.8: Depression rates are not associated with year over year population change. Results from ordinary least squares fits with the rate of population change included. (Continued)

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
NSDUH Ln Population	-0.0952	0.045	-2.108	0.044	-0.188	-0.003
NSDUH Population Change %	67.5237	130.243	0.518	0.608	-199.266	334.314

7.2 Appendix B

Supplementary Text

Derivation of the homophily and hetrophobia adjustments

Here we expand on the derivation of the homophily and heterophobia adjustments in the main text. We start with Equation 3.2 of the main text which gives the average number of social interactions per individual in group g. From this we can write down the average number of social interactions per individual for the entire city as a sum over the different groups, G. For simplicity we have dropped the residual term e^{ξ_i}:

$$k_i \sim \frac{a_0 l}{A_i N_i}[\sum_{g=1}^{G}(N_{g,i}(1+h_{g,i}^{hom}) + \sum_{j \neq g} N_{j,i}(1-h_{g,i}^{het})) \cdot N_{g,i}] \tag{7.1}$$

multiplying through by $N_{g,i}$ we have:

$$k_i \sim \frac{a_0 l}{A_i N_i}[\sum_{g=1}^{G}(N_{g,i}^2) + \sum_{g=1}^{G}(N_{g,i}^2 h_{g,i}^{hom}) + \sum_{g=1:G}\sum_{j\neq g} N_{g,i}N_{g,j} - \sum_{g=1:G}\sum_{j\neq g} N_{g,i}N_{j,i}h_{g,i}^{het}] \quad (7.2)$$

since the third term in the brackets gives two copies of each $N_{g,i}N_{j,i}$ term, we can then write:

$$k_i \sim \frac{a_0 l}{A_i N_i}[(\sum_{g=1}^{G} N_{g,i})^2 - \sum_{g=1}^{G}\sum_{j=g+1}^{G} N_{g,i}N_{j,i}(h_{g,i}^{het} + h_{j,i}^{het}) + \sum_{g=1}^{G}(N_{g,i}^2 h_{g,i}^{hom})] \quad (7.3)$$

and finally, we divide and multiply the second term by N_i^2 and arrive back at Equation 3.3 of the main text:

$$k_i \sim \frac{a_0 l}{A_i N_i}[N_i^2 - N_i^2 \sum_{g=1}^{G}\sum_{j=g+1}^{G} \frac{N_{g,i}}{N_i}\frac{N_{j,i}}{N_i}(h_{g,i}^{het} + h_{j,i}^{het}) + N_i^2 \sum_{g=1}^{G}(\frac{N_{g,i}}{N_i})^2 h_{g,i}^{hom}] \quad (7.4)$$

equivalently:

$$k_i \sim \frac{a_0 l N_i}{A_i}[1 - A_i^{het} + A_i^{hom}] \quad (7.5)$$

Note that there are no homophily and heterophobia effects Equation 7.4 becomes:

$$k_i \sim \frac{a_0 l}{A_i N_i}[N_i^2] = \frac{a_0 l N_i}{A_i} \quad (7.6)$$

and we recover the typical scaling law.

In the main text, we made a comment that A_i^{het} is always less than 0.5 and ≥ 0. This can be easily seen from the fact that A_i^{het} is at its smallest when there are only two groups and those groups are equally balanced. In this case, A_i^{het} takes on a value of $0.5 \cdot 0.5 \cdot (1+1)$ when the two groups are completely heterophobic and is smaller for groups with less heterophobia, imbalanced proportions of the population, or for more than two groups. In general, for $G > 2$ groups, A_i^{het} is at most $2/G^2 \cdot \binom{G}{2} = (G-1)/G$ which occurs when the heterophobia values

for all groups are 1 and groups are equal in size. Similarly, A_i^{hom} is at minimum $1/G$ and this occurs when homophily values for all groups are 1 and groups are equal in size.

As a note, in general, heterophobia values need not be the same across all groups. In particular, we can define a city specific matrix with entries $\in [-1, 1]$ which specifies the degree to which individuals avoid or preferentially interact with individuals from the same or other groups:

$$H_i = \begin{bmatrix} h_{11i} & h_{12i} & \cdots & h_{1Gi} \\ h_{21i} & h_{22i} & \cdots & h_{2Gi} \\ \vdots & \vdots & \ddots & \vdots \\ h_{G1i} & h_{G2i} & \cdots & h_{GGi} \end{bmatrix} \tag{7.7}$$

The relative rate of interactions between any two groups (or within a group) is given by $1 + H_i$. In the main text we assumed homophily and heterophobia, i.e., that the diagonal elements of H are positive and the off diagonal elements of H are negative. However, in general, homophobia and heterophilly may also be present so that the entries of H can take on positive or negative values. In this case, positive values correspond to homophily or heterophily and negative values correspond to homophobia or heterophobia.

Our specific constraints on H in the models presented in the main text were that: (1) homophily is specified by the diagonal entries of H_i, $h_{g,i}^{hom} = h_{ggi}$, which we assume to be positive, and (2) that each group avoids all other groups equally, so that there are repeated off diagonal entries and that all of the off diagonal entries are negative. Specifically, that $h_{g,i}^{het} = h_{gji}$ for all $j \neq g$.

This notation allows for an alternative notation for Equation 7.4:

$$k_i \sim \frac{a_0 l}{A_i N_i}[tr(\mathbf{N_i} H_i \mathbf{N_i}) + tr([1 - \mathbf{I}]\mathbf{N_i} H_i \mathbf{N_i}] \tag{7.8}$$

where $\mathbf{N_i}$ is the diagonal matrix of group sizes and I is the identity matrix. In other words, between-group interactions are given by the sum of off diagonal elements of $\mathbf{N_i} H \mathbf{N_i}$ and

within-group interactions are given by the sum of diagonal elements.

Census Block Group Analysis

All of the analyses described in the main text were also conducted with census block groups instead of census tracts. In contrast to census tracts which contain, on average, 4,000 individuals, census block groups contain 1,500 individuals on average. A similar pattern of results was found for census block groups.

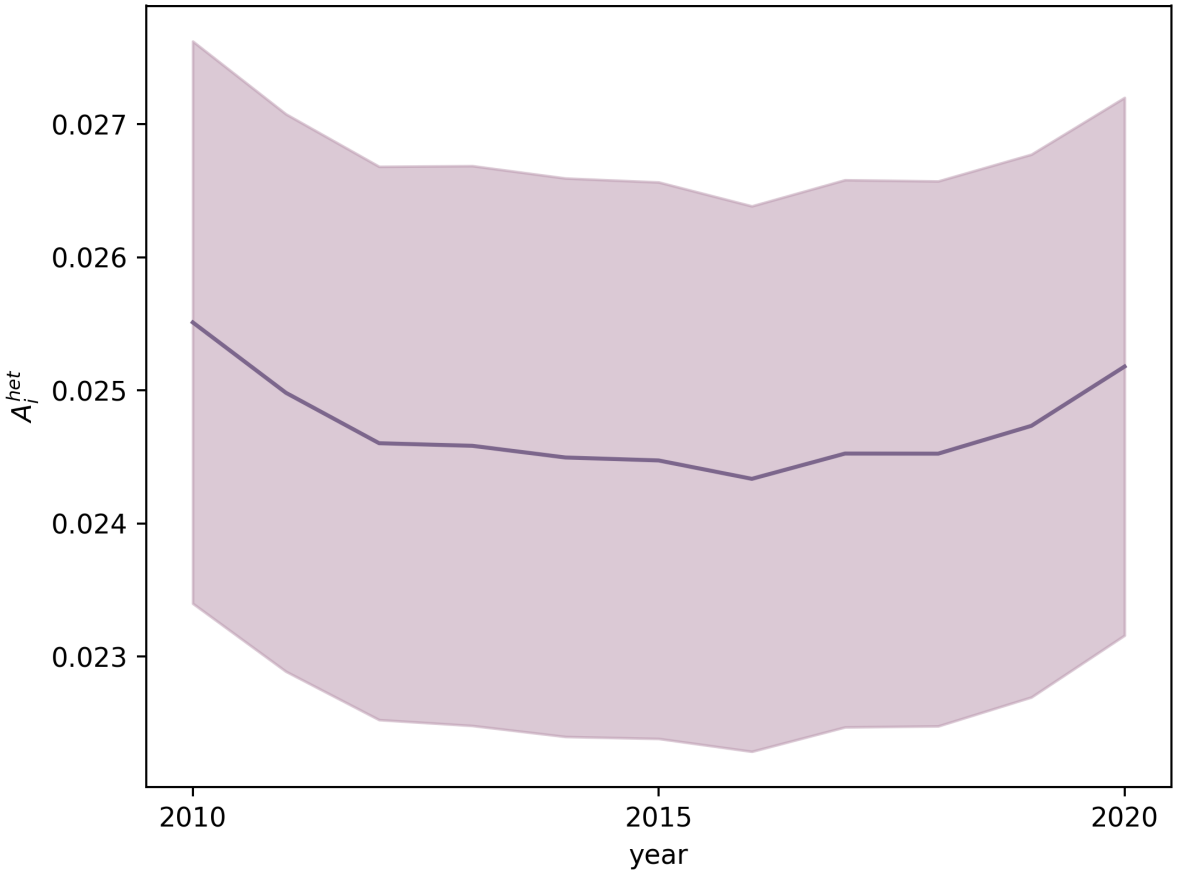
Table 7.9: Fits of calculated heterophobia adjustments to median income scaling deviations by year. b^{het} determines the strength of the coupling between economic productivity and residential segregation by controlling levels of heterophobia associated with residential segregation.

year	b^{het} 95% CI	R^2
2013	[1.03, 1.60]	0.09
2014	[1.10, 1.67]	0.10
2015	[1.25, 1.84]	0.11
2016	[1.35, 1.93]	0.13
2017	[1.39, 1.96]	0.13
2018	[1.37, 1.96]	0.13
2019	[1.46, 2.04]	0.14
2020	[1.46, 2.04]	0.14

Table 7.10: Fits of calculated heterophobia adjustments to GDP scaling deviations by year.

year	b^{het} 95% CI	R^2
2013	[-0.12, 0.93]	0.00283
2014	[0.01, 1.06]	0.00476
2015	[0.07, 1.11]	0.00597
2016	[0.17, 1.18]	0.00807
2017	[0.19, 1.20]	0.00873
2018	[0.18, 1.19]	0.00831
2019	[0.27, 1.23]	0.01072
2020	[0.26, 1.21]	0.01076

Supplementary Figures



Supplementary Figure 7.1: Mean A_i^{het} across cities over time. The shaded region represents the 95% interval of the standard error of the mean.

Supplementary Tables

Supplementary Table 7.11: Spearman Rank Order Correlation between Median Income and GDP by year.

year	r_s	p-value
2010	0.56	1.77e-74
2011	0.56	3.38e-72
2012	0.55	2.52e-71
2013	0.54	2.08e-66
2014	0.53	6.96e-64
2015	0.52	1.19e-60
2016	0.52	1.34e-62
2017	0.53	7.85e-65
2018	0.54	1.97e-67
2019	0.55	8.83e-71
2020	0.56	5.50e-73

Supplementary Table 7.12: Fits of calculated heterophobia corrections to median income scaling deviations by year.

year	b^{het} 95% CI	R^2	n
2010	[-1.88, -1.23]	0.09	872
2011	[-1.94, -1.28]	0.10	872
2012	[-2.02, -1.34]	0.10	873
2013	[-1.96, -1.28]	0.09	851
2014	[-2.06, -1.37]	0.10	852
2015	[-2.20, -1.49]	0.11	852
2016	[-2.28, -1.56]	0.11	865
2017	[-2.31, -1.60]	0.12	865
2018	[-2.31, -1.58]	0.11	866
2019	[-2.44, -1.71]	0.13	874
2020	[-2.49, -1.77]	0.13	874

Supplementary Table 7.13: Fits of calculated homophily and heterophobia corrections to median income scaling deviations by year.

year	b^{het} 95% CI	b^{hom} 95% CI	R^2	n
2010	[-2.19, -1.15]	[-0.35, 0.63]	0.09	872
2011	[-2.20, -1.14]	[-0.43, 0.57]	0.10	872
2012	[-2.37, -1.29]	[-0.33, 0.69]	0.10	873
2013	[-2.33, -1.25]	[-0.31, 0.71]	0.09	851
2014	[-2.41, -1.33]	[-0.32, 0.71]	0.10	852
2015	[-2.53, -1.42]	[-0.37, 0.69]	0.11	852
2016	[-2.69, -1.58]	[-0.26, 0.79]	0.11	865
2017	[-2.59, -1.48]	[-0.43, 0.63]	0.12	865
2018	[-2.49, -1.35]	[-0.58, 0.51]	0.11	866
2019	[-2.52, -1.40]	[-0.70, 0.41]	0.13	874
2020	[-2.36, -1.28]	[-1.00, 0.14]	0.14	874

Supplementary Table 7.14: Fits of calculated heterophobia corrections to GDP scaling deviations by year.

year	b^{het} 95% CI	R^2	n
2010	[-0.94, 0.20]	0.00191	861
2011	[-0.98, 0.23]	0.00172	863
2012	[-1.10, 0.13]	0.00280	862
2013	[-1.04, 0.17]	0.00237	843
2014	[-1.26, -0.02]	0.00488	844
2015	[-1.38, -0.17]	0.00738	847
2016	[-1.44, -0.25]	0.00910	860
2017	[-1.39, -0.20]	0.00798	859
2018	[-1.38, -0.18]	0.00762	858
2019	[-1.48, -0.34]	0.01135	865
2020	[-1.51, -0.40]	0.01303	868

Supplementary Table 7.15: Fits of calculated homophily and heterophobia corrections to median income scaling deviations by year using the segregation index.

year	b^{het} 95% CI	b^{hom} 95% CI	R^2	n
2010	[-0.76, -0.45]	[-0.06, 0.17]	0.07	872
2011	[-0.79, -0.47]	[-0.07, 0.17]	0.07	872
2012	[-0.82, -0.50]	[-0.06, 0.19]	0.07	873
2013	[-0.79, -0.46]	[-0.07, 0.18]	0.06	851
2014	[-0.84, -0.51]	[-0.08, 0.17]	0.07	852
2015	[-0.90, -0.56]	[-0.10, 0.15]	0.08	852
2016	[-0.94, -0.59]	[-0.04, 0.23]	0.08	865
2017	[-0.96, -0.62]	[-0.09, 0.18]	0.08	865
2018	[-0.96, -0.61]	[-0.14, 0.14]	0.08	866
2019	[-1.02, -0.67]	[-0.17, 0.11]	0.09	874
2020	[-1.05, -0.70]	[-0.24, 0.04]	0.10	874

Supplementary Table 7.16: Fits of calculated heterophobia corrections to GDP scaling deviations by year using the segregation index.

year	π 95% CI	R^2	n
2010	[-0.30, 0.24]	0.00006	861
2011	[-0.32, 0.25]	0.00007	863
2012	[-0.37, 0.20]	0.00039	862
2013	[-0.34, 0.23]	0.00015	843
2014	[-0.46, 0.13]	0.00144	844
2015	[-0.55, 0.03]	0.00366	847
2016	[-0.58, -0.02]	0.00508	860
2017	[-0.57, -0.00]	0.00458	859
2018	[-0.55, 0.02]	0.00375	858
2019	[-0.57, -0.03]	0.00533	865
2020	[-0.60, -0.07]	0.00706	868

Supplementary Table 7.17: Fits of calculated homophily and heterophobia corrections to median income scaling deviations by year using the gini coefficient.

year	b^{het} 95% CI	b^{hom} 95% CI	R^2	n
2010	[-0.59, -0.35]	[-0.06, 0.12]	0.06	872
2011	[-0.61, -0.37]	[-0.08, 0.11]	0.07	872
2012	[-0.63, -0.38]	[-0.06, 0.13]	0.07	873
2013	[-0.61, -0.36]	[-0.07, 0.12]	0.06	851
2014	[-0.64, -0.39]	[-0.08, 0.13]	0.07	852
2015	[-0.69, -0.43]	[-0.10, 0.11]	0.07	852
2016	[-0.72, -0.45]	[-0.05, 0.16]	0.08	865
2017	[-0.74, -0.47]	[-0.08, 0.13]	0.08	865
2018	[-0.74, -0.47]	[-0.14, 0.08]	0.08	866
2019	[-0.79, -0.52]	[-0.16, 0.06]	0.09	874
2020	[-0.81, -0.55]	[-0.23, -0.00]	0.11	874

Supplementary Table 7.18: Fits of calculated heterophobia corrections to GDP scaling deviations by year using the gini coefficient.

year	π 95% CI	R^2	n
2010	[-0.22, 0.20]	0.00001	861
2011	[-0.24, 0.20]	0.00004	863
2012	[-0.28, 0.17]	0.00029	862
2013	[-0.25, 0.19]	0.00007	843
2014	[-0.34, 0.12]	0.00111	844
2015	[-0.40, 0.05]	0.00279	847
2016	[-0.43, 0.01]	0.00425	860
2017	[-0.42, 0.02]	0.00380	859
2018	[-0.41, 0.03]	0.00335	858
2019	[-0.43, -0.01]	0.00473	865
2020	[-0.45, -0.05]	0.00692	868

Supplementary Table 7.19: Fits of calculated homophily and heterophobia corrections to median income scaling deviations by year using the exposure index.

year	b^{het} 95% CI	b^{hom} 95% CI	R^2	n
2010	[-1.43, -0.70]	[-0.09, 0.44]	0.06	872
2011	[-1.47, -0.72]	[-0.12, 0.42]	0.07	872
2012	[-1.57, -0.80]	[-0.08, 0.48]	0.07	873
2013	[-1.53, -0.76]	[-0.08, 0.48]	0.06	851
2014	[-1.60, -0.82]	[-0.08, 0.49]	0.07	852
2015	[-1.72, -0.91]	[-0.06, 0.52]	0.07	852
2016	[-1.81, -0.99]	[-0.02, 0.58]	0.08	865
2017	[-1.78, -0.96]	[-0.10, 0.51]	0.08	865
2018	[-1.73, -0.87]	[-0.22, 0.43]	0.08	866
2019	[-1.81, -0.95]	[-0.26, 0.40]	0.09	874
2020	[-1.65, -0.84]	[-0.45, 0.21]	0.11	874

Supplementary Table 7.20: Fits of calculated heterophobia corrections to GDP scaling deviations by year using the exposure index.

year	π 95% CI	R^2	n
2010	[-0.54, 0.26]	0.00056	861
2011	[-0.54, 0.32]	0.00028	863
2012	[-0.61, 0.26]	0.00073	862
2013	[-0.59, 0.27]	0.00063	843
2014	[-0.74, 0.14]	0.00212	844
2015	[-0.88, -0.01]	0.00471	847
2016	[-0.91, -0.05]	0.00565	860
2017	[-0.88, -0.02]	0.00487	859
2018	[-0.87, 0.01]	0.00434	858
2019	[-0.94, -0.10]	0.00683	865
2020	[-1.01, -0.20]	0.00992	868

Supplementary Table 7.21: Fits of calculated heterophobia adjustments to GDP scaling deviations by year with outliers included.

year	b^{het} 95% CI	R^2	n
2010	[-1.28, 0.04]	0.00389	875
2011	[-1.32, 0.07]	0.00355	875
2012	[-1.50, -0.07]	0.00528	875
2013	[-1.51, -0.09]	0.00570	855
2014	[-1.72, -0.27]	0.00834	855
2015	[-1.72, -0.37]	0.01071	855
2016	[-1.65, -0.34]	0.01011	868
2017	[-1.63, -0.28]	0.00885	868
2018	[-1.67, -0.27]	0.00847	868
2019	[-1.69, -0.36]	0.01027	875
2020	[-1.60, -0.38]	0.01147	875

Supplementary Table 7.22: Fits of calculated heterophobia adjustments to median income scaling deviations by year with outliers included.

year	b^{het} 95% CI	R^2	n
2010	[-2.50, -1.19]	0.06	885
2011	[-2.49, -1.16]	0.06	885
2012	[-2.58, -1.26]	0.06	885
2013	[-2.58, -1.27]	0.07	864
2014	[-2.76, -1.44]	0.07	864
2015	[-2.95, -1.61]	0.08	864
2016	[-3.23, -1.90]	0.09	877
2017	[-3.27, -1.94]	0.10	877
2018	[-3.39, -2.06]	0.10	876
2019	[-3.54, -2.20]	0.11	885
2020	[-3.42, -2.12]	0.11	885

7.3 Appendix C

Supplementary Text

Here we reproduce the parts of our previously developed extension to the standard Urban Scaling Theory model Stier et al. [2022b] that are relevant for inter-group interactions and implicit racial biases.

The extension on which this manuscript is based starts from the standard formulation of Urban scaling theory Bettencourt [2013, 2021b], which describes cities as spatially embedded networks of socioeconomic interactions.

For the average per-capita number of social interactions, k, the urban scaling law takes form of $k \sim N^{\delta}$, where $\delta = \frac{1}{6}$. This form of the scaling law results from a mean-field approximation that individuals interact homogeneously. Under these conditions, we take individuals to have an interaction cross section a_0 and a characteristic travel length l per unit time. This gives the average number of interactions for a large city ($N >> 1$) as :

$$k \sim \frac{a_0 l}{A_n} N \tag{7.1}$$

, where A_n is the area of the city's networks. The scaling law for the area of the city's networks, $A_n \sim N^{1-\delta}$ Bettencourt [2013], recovers the scaling law for k.

In order to derive the model for heterogeneous group interactions it is important to understand the standard formulation of the scaling law for interactions. We start by observing that, $\frac{a_0 l}{A_n}$ is a probability composed of the fraction of a city's area that individuals cover, on average, over a given time period. This is the average probability of interacting with other individuals in the city. Thus, the total expected number of interactions for each individual is given by their probability of interacting, $\frac{a_0 l}{A_n}$, multiplied by the number of individuals they could interact with, N.

Next, we extend this standard formulation by modeling each individual in the city as

belonging to some distinct number of groups, indexed by g. Individuals in these groups may interact with a lower probability with other groups. We define this relative reduction in out-group interactions by $1 - h_g^{het}$, where $h_g^{het} \in [0, 1]$ is the heterophobia of group g.

With these definitions, the average number of interactions for individuals in group g with individuals in different groups is given by:

$$k_{g,inter} \sim \frac{a_0 l}{A} \sum_{j \neq g} N_j (1 - h_g^{het}) \tag{7.2}$$

where N_g is the population of focal group g. The total number of between-group social interactions for all individuals in group g is $k_{g,inter} N_g$, on average. Therefore, the average number of between-group social interactions for individuals in an observed segregated city i with G different groups, is $k_i \sim \frac{1}{N} \sum_{g=1}^{G} k_{g,inter,i} N_{g,i}$.

This brings us almost to Equation 2 of the main text:

$$k_{i,inter} \sim \frac{a_0 l}{A_i N_i} \sum_{j \neq g} N_{j,i} (1 - h_{g,i}^{het})) \cdot N_{g,i} \tag{7.3}$$

Multiplying and dividing through by N_i^2 we have:

$$k_{i,inter} \sim \frac{a_0 l N_i^2}{A_i N_i} \sum_{j \neq g} \frac{N_{j,i}}{N_i} \frac{N_{g,i}}{N_i} (1 - h_{g,i}^{het}))\cdot \tag{7.4}$$

which directly simplifies to the between-group interactions terms of Equation 2 of the main text:

$$k_{i,inter} \sim N^{\delta} \sum_{j \neq g} \frac{N_{j,i}}{N_i} \frac{N_{g,i}}{N_i} (1 - h_{g,i}^{het}))\cdot \tag{7.5}$$

since $\frac{a_0 l N_i}{A_i} \sim N^{\delta}$.

Note that when heterophobia is complete, i.e., $h_{g,i}^{het} = 0$ or there is only one group in the city, $k_{i,inter}$ goes to 0, as expected, and there are no between-group interactions.

As a note, in general, heterophobia values need not be the same across all groups and we can define a matrix with entries $\in [-1, 1]$ which specifies the degree to which individuals avoid or preferentially interact with individuals from the same or other groups:

$$H_i = \begin{bmatrix} h_{11i} & h_{12i} & \cdots & h_{1Gi} \\ h_{21i} & h_{22i} & \cdots & h_{2Gi} \\ \vdots & \vdots & \ddots & \vdots \\ h_{G1i} & h_{G2i} & \cdots & h_{GGi} \end{bmatrix} \tag{7.6}$$

The relative rate of interactions between any two groups (or within a group) is given by $1+H_i$. In the model presented here, we constrain H so that each group avoids all other groups equally and that there are repeated off-diagonal entries that are all negative. Specifically, that $h^{het}_{g,i} = h_{gji}$ for all $j \neq g$.

This notation allows for an alternative notation for Equation 5:

$$k_{i,inter} \sim \frac{a_0 l}{A_i N_i} tr\left([1 - \mathbf{I}]\mathbf{N_i} H_i \mathbf{N_i}\right) \tag{7.7}$$

where $\mathbf{N_i}$ is the diagonal matrix of group sizes and I is the identity matrix. In other words, between-group interactions are given by the sum of off-diagonal elements of $\mathbf{N_i} H \mathbf{N_i}$.

Inter-group interactions with two groups

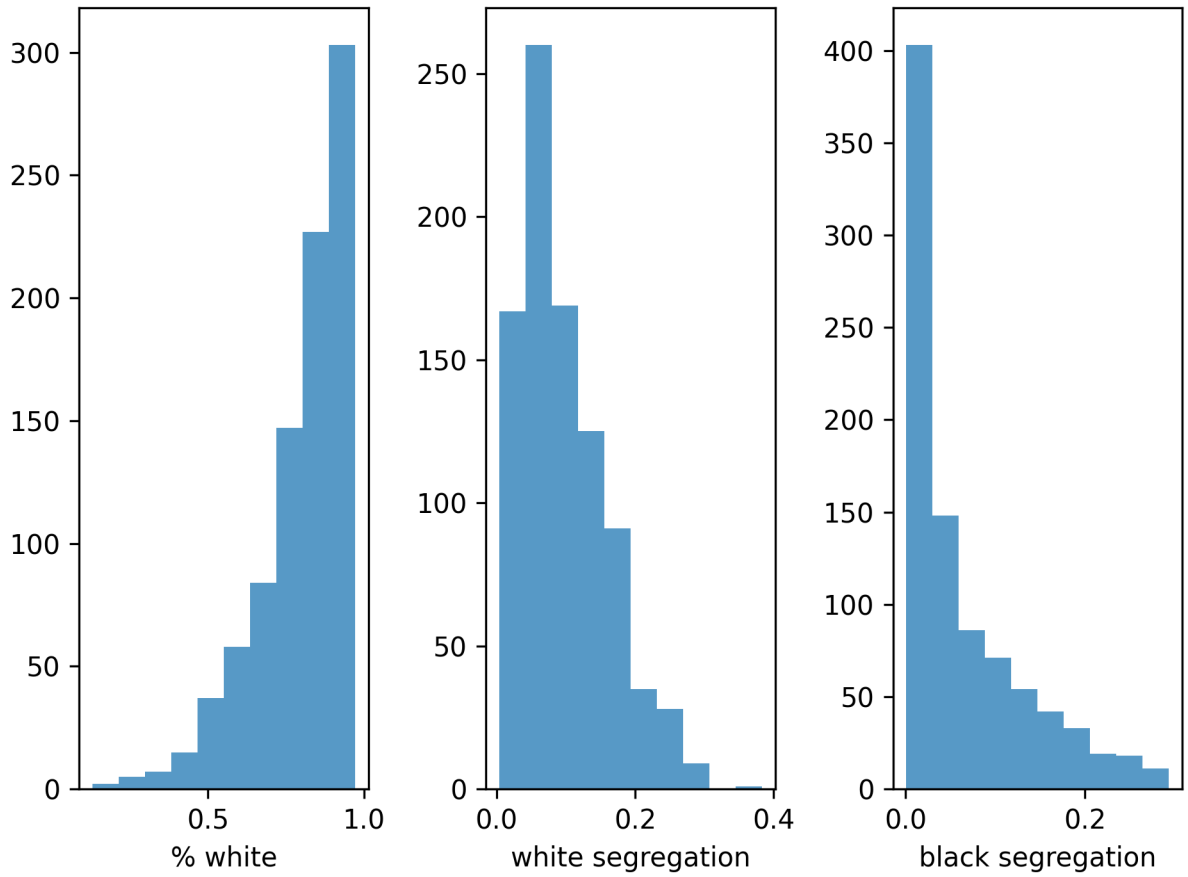
In the case of two groups, the inter-group term of Equation 3 of the main text becomes:

$$k_{i,inter} \sim N_i^{\delta}[2 \cdot \frac{N_{1,i}}{N_i}\frac{N_{2,i}}{N_i} - \frac{N_{1,i}}{N_i}\frac{N_{2,i}}{N_i} h^{het}_{1,i} - \frac{N_{1,i}}{N_i}\frac{N_{2,i}}{N_i} h^{het}_{2,i}] \tag{7.8}$$

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

$$k_{i,inter} \sim N^{\delta}[(\frac{N_{1,i}}{N_i}\frac{N_i - N_{1,i}}{N_i}) \cdot (2 - h^{het}_{1,i} - h^{het}_{2,i})] \tag{7.9}$$

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 Figure 7.1: 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 7.23: 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	$\mathbf{s_1} + \mathbf{s_2}$ β_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 Figures

Supplementary Table 7.24: 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 7.25: 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 4.3 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
Maj Grp Sz Adj	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 7.26: 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 4.3 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> |z|**	**[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
Maj Grp Sz Adj	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 7.27: 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 4.3 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
Maj Grp Sz Adj	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 7.28: 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 4.3 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
Maj Grp Sz Adj	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 7.29: 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 4.3 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
Maj Grp Sz Adj	-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 7.30: 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 4.3 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
Maj Grp Sz Adj	-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 7.31: 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 4.3 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
Maj Grp Sz Adj	-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 7.32: 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 4.3 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
Maj Grp Sz Adj	-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 7.33: 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 4.3 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
Maj Grp Sz Adj	-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 7.34: 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 4.3 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
Maj Grp Sz Adj	-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 7.35: 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 4.3 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
Maj Grp Sz Adj	-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 7.36: 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 7.37: 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 7.38: 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 7.39: 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 7.40: 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 7.41: 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 7.42: 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 7.43: 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 7.44: 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 7.45: 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 7.46: 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 7.47: 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 7.48: 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 7.49: 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 7.50: 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 7.51: 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 7.52: 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 7.53: 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 7.54: 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 7.55: 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 7.56: 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 7.57: 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 7.58: 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 7.59: 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 7.60: 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 7.61: 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 7.62: 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 7.63: 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 7.64: 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.

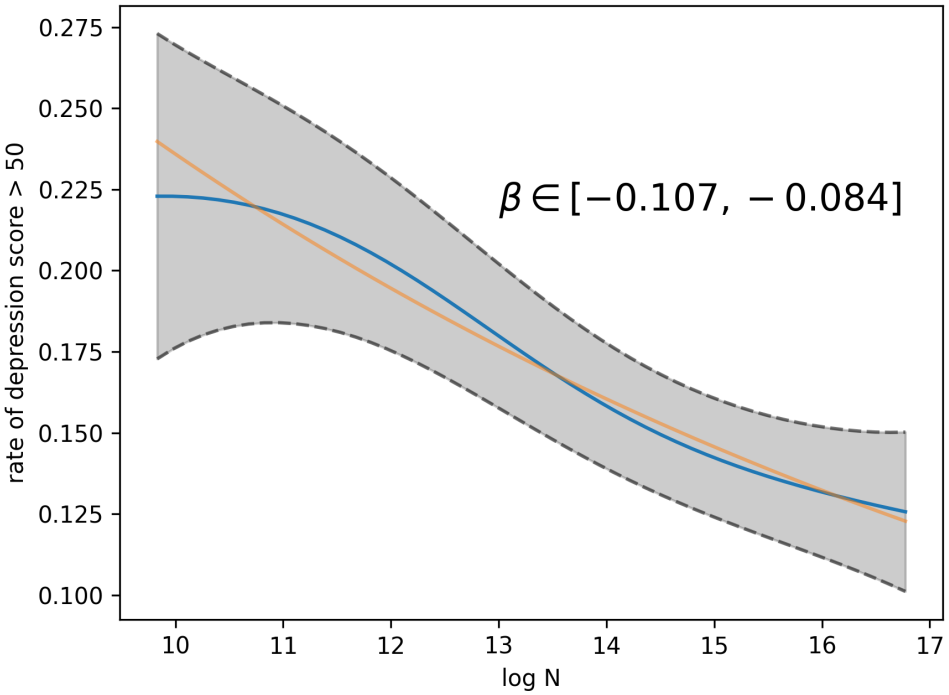
year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.053	0.143	0.040	0.288	31
2011	0.112	0.145	0.004	0.307	31
2012	0.047	0.082	0.121	0.312	26
2013	0.052	0.195	0.136	0.418	28
2014	0.039	0.262	0.079	0.408	34
2015	0.014	0.340	0.134	0.508	43
2016	0.115	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 7.65: 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.

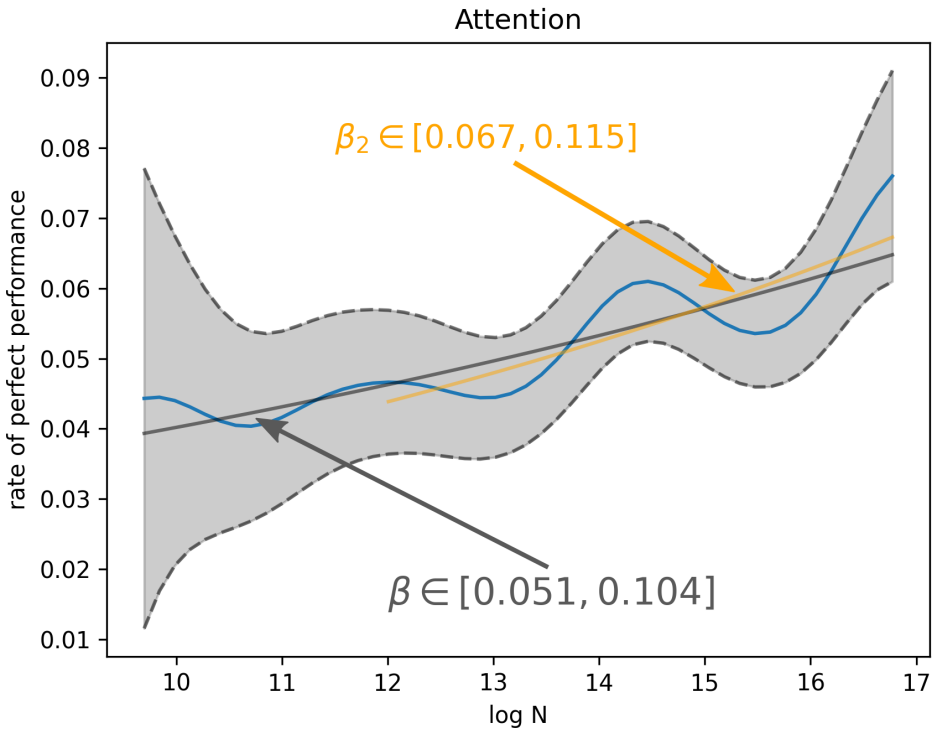
year	scaling R^2	majority group size adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.125	0.128	0.031	0.281	126
2011	0.056	0.120	0.027	0.212	119
2012	0.165	0.047	0.007	0.231	88
2013	0.134	0.091	0.059	0.282	98
2014	0.097	0.066	0.009	0.184	116
2015	0.044	0.158	0.050	0.258	129
2016	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 7.66: 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.

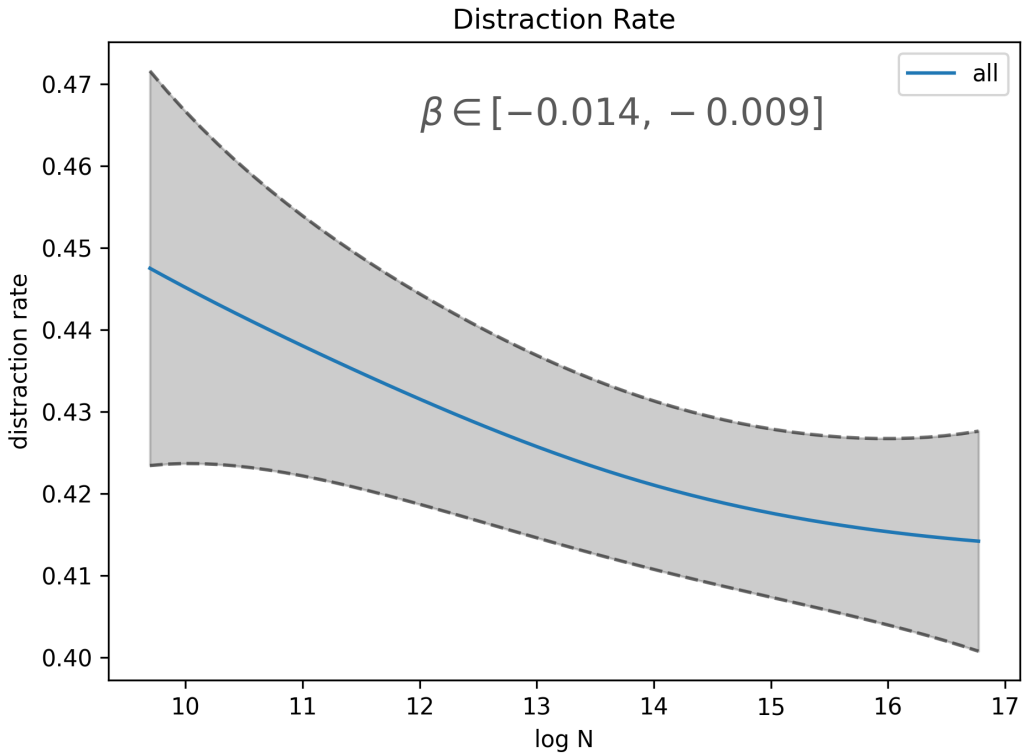
year	scaling R^2	Majority Group Size Adjustment R^2	heterophobia adjustment R^2	overall R^2	# cities
2010	0.053	0.143	0.044	0.291	31
2011	0.112	0.145	0.010	0.310	31
2012	0.047	0.082	0.120	0.311	26
2013	0.052	0.195	0.128	0.410	28
2014	0.039	0.262	0.079	0.408	34
2015	0.014	0.340	0.118	0.494	43
2016	0.115	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



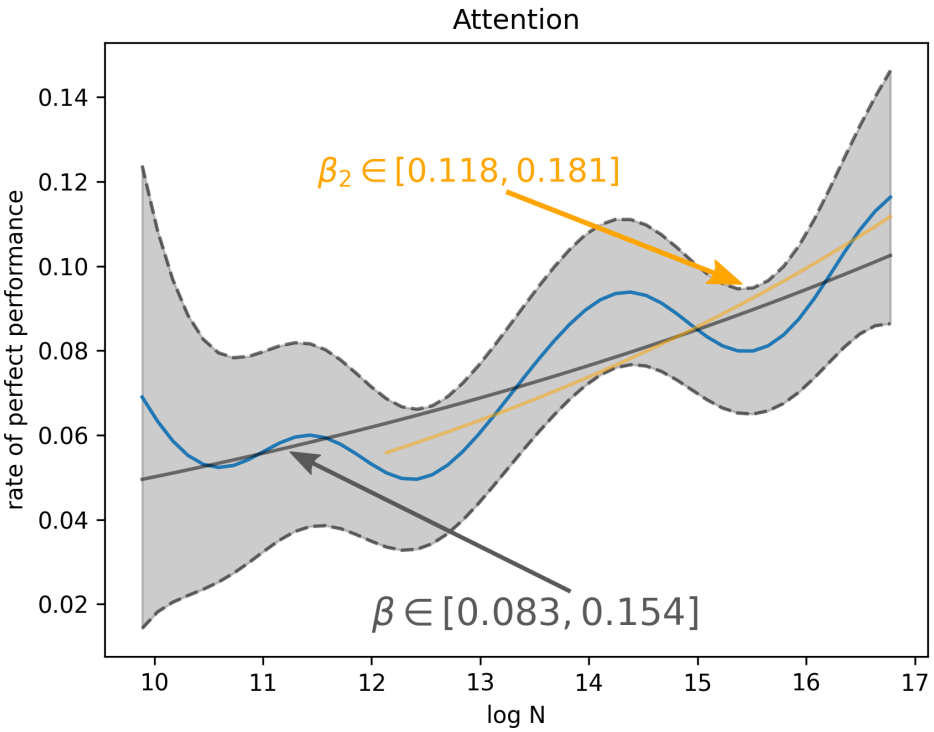
Supplementary Figure 7.1: Depression rates are lower in larger cities. The rate of individuals with high depression scores is approximately $\beta = 0.085$. The blue line shows the Nadaraya–Watson kernel regression estimate and the envelope with dashed lines shows the 95% confidence interval for the kernel regression.



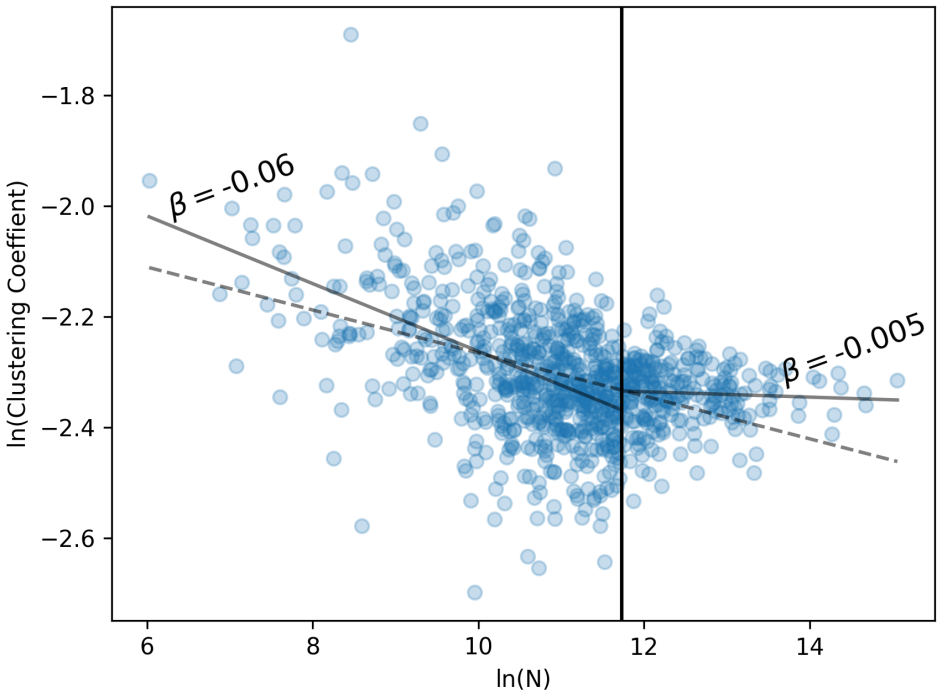
Supplementary Figure 7.2: Attention increases in larger cities even when including participants who report they were distracted during the task.



Supplementary Figure 7.3: The rate of participants who were distracted during the attention task decreased slightly with city size. The blue line shows the Nadaraya–Watson kernel regression estimate and the envelope with dashed lines shows the 95% confidence interval for the kernel regression.



Supplementary Figure 7.4: Attention increases in larger cities even when only including data from participants' first time completing the attention task.



Supplementary Figure 7.5: Social Network Clustering decreases with city population size. In data from a large online social network, average local clustering

Supplementary Tables

Supplementary Table 7.1: Scaling slope estimates and upper and lower 95% confidence bounds for different thresholds for depression.

threshold	beta	lower	upper
40	-0.081897	-0.089391	-0.073341
41	-0.086149	-0.094957	-0.077788
42	-0.091519	-0.114251	-0.062688
43	-0.081103	-0.090140	-0.071892
44	-0.084772	-0.093766	-0.075670
45	-0.088288	-0.098103	-0.077953
46	-0.084351	-0.095333	-0.072959
47	-0.083739	-0.094813	-0.072771
48	-0.086673	-0.098015	-0.075639
49	-0.095199	-0.106052	-0.083338
50	-0.096373	-0.106973	-0.084122
51	-0.087315	-0.098771	-0.075796
52	-0.075026	-0.085952	-0.063509
53	-0.085681	-0.098720	-0.072464
54	-0.084602	-0.097370	-0.070910
55	-0.092216	-0.106983	-0.076881

Supplementary Table 7.2: Logistic regression for hit rate > 0.9 including city population and individual dempographics.

Dep. Variable:	hit		**No. Observations:**		3082	
Model:	Logit		**Df Residuals:**		3053	
Method:	MLE		**Df Model:**		28	
Date:	Thu, 09 Feb 2023		**Pseudo R-squ.:**		0.05018	
Time:	15:20:16		**Log-Likelihood:**		-748.03	
converged:	True		**LL-Null:**		-787.55	
Covariance Type:	nonrobust		**LLR p-value:**		9.283e-07	
	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	-3.0406	1.129	-2.693	0.007	-5.253	-0.828
log population	**0.1194**	**0.047**	**2.546**	**0.011**	**0.027**	**0.211**
age	-0.3018	0.158	-1.912	0.056	-0.611	0.008
female	**-0.0408**	**0.006**	**-6.607**	**0.000**	**-0.053**	**-0.029**
white	0.0244	0.227	0.108	0.914	-0.420	0.469
black	0.0194	0.299	0.065	0.948	-0.567	0.606
asian	0.1518	0.327	0.464	0.643	-0.490	0.793
latino	0.1635	0.210	0.779	0.436	-0.248	0.575
indian	0.0565	0.525	0.108	0.914	-0.972	1.085
pacific islander	-0.9949	1.038	-0.958	0.338	-3.029	1.040
native	0.4016	0.340	1.180	0.238	-0.266	1.069
high school/GED	0.5811	0.769	0.755	0.450	-0.927	2.089
some college	0.5928	0.755	0.785	0.432	-0.887	2.072
2year	0.6441	0.771	0.835	0.404	-0.867	2.155
4year	0.6448	0.755	0.854	0.393	-0.834	2.124
graduate school	0.8445	0.758	1.114	0.265	-0.642	2.331
income 1	0.1534	0.350	0.438	0.661	-0.533	0.840
income 2	-0.1627	0.364	-0.447	0.655	-0.876	0.550
income 3	-0.1957	0.356	-0.549	0.583	-0.894	0.503
income 4	-0.0926	0.340	-0.273	0.785	-0.759	0.574
income 5	-0.0376	0.341	-0.110	0.912	-0.707	0.631
income 6	-0.2353	0.351	-0.670	0.503	-0.924	0.453
income 7	-0.1415	0.361	-0.392	0.695	-0.849	0.566
income 8	-0.7203	0.515	-1.398	0.162	-1.730	0.289
income 9	0.1574	0.560	0.281	0.779	-0.941	1.256
fair health	-0.4194	0.481	-0.872	0.383	-1.362	0.524
good health	0.1324	0.450	0.294	0.768	-0.749	1.014
very good health	0.1628	0.458	0.355	0.722	-0.735	1.061
excellent health	-0.0714	0.537	-0.133	0.894	-1.123	0.980

Supplementary Table 7.3: Logistic regression for hit rate > 0.9 including stress and sleep variables and city population.

Dep. Variable:	hit	**No. Observations:**	2224
Model:	Logit	**Df Residuals:**	2216
Method:	MLE	**Df Model:**	7
Date:	Thu, 13 Apr 2023	**Pseudo R-squ.:**	0.01177
Time:	10:02:14	**Log-Likelihood:**	-603.18
converged:	True	**LL-Null:**	-610.36
Covariance Type:	nonrobust	**LLR p-value:**	0.04508

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	-5.2671	1.918	-2.747	0.006	-9.026	-1.508
log population	**0.1341**	**.050**	**2.656**	**0.008**	**0.035**	**0.233**
short term stress	0.4859	0.901	0.539	0.590	-1.280	2.252
short term stress squared	-0.0014	0.130	-0.011	0.991	-0.256	0.253
chronic stress	0.7754	0.546	1.420	0.156	-0.295	1.846
chronic stress squared	0.0298	0.089	0.335	0.738	-0.145	0.204
short term x chronic stress	-0.2382	0.125	-1.913	0.056	-0.482	0.006
morning sleep quality	-0.2099	0.203	-1.036	0.300	-0.607	0.187

Supplementary Table 7.4: Sensitivity test for the threshold for poor performance with all data included. Note that the negative coefficient indicates that there are fewer poor performers in larger cities.

hit rate threshold	β	low	high
0.450	-0.033	-0.041	-0.026
0.460	-0.033	-0.041	-0.025
0.470	-0.033	-0.040	-0.025
0.480	-0.033	-0.041	-0.025
0.490	-0.033	-0.040	-0.025
0.500	-0.033	-0.040	-0.025
0.510	-0.014	-0.018	-0.010
0.520	-0.014	-0.018	-0.010
0.530	-0.014	-0.018	-0.010
0.540	-0.014	-0.018	-0.010
0.550	-0.013	-0.017	-0.009

Supplementary Table 7.5: Sensitivity test for the threshold for low false alarm rate. Note that the positive coefficient indicates that there are more high performers in larger cities.

false alarm threshold	beta	low	high
0.150	0.000	-0.004	0.004
0.160	0.000	-0.004	0.003
0.170	0.010	0.007	0.013
0.180	0.010	0.007	0.014
0.190	0.010	0.007	0.014
0.200	0.010	0.007	0.013
0.210	0.009	0.006	0.012
0.220	0.009	0.006	0.012
0.230	0.010	0.007	0.013
0.240	0.011	0.007	0.014
0.250	0.011	0.008	0.014

Supplementary Table 7.6: Sensitivity test for the threshold for high false alarm rate. Note that the negative coefficient indicates that there are fewer poor performers in larger cities.

false alarm threshold	beta	low	high
0.850	0.000	-0.085	0.073
0.840	0.000	-0.068	0.071
0.830	0.000	-0.018	0.017
0.820	0.000	-0.018	0.018
0.810	-0.009	-0.011	-0.006
0.800	-0.009	-0.011	-0.006
0.790	-0.009	-0.011	-0.006
0.780	-0.009	-0.011	-0.006
0.770	-0.009	-0.011	-0.006
0.760	-0.009	-0.012	-0.006
0.750	-0.009	-0.012	-0.006

Supplementary Table 7.7: Sensitivity test for the threshold for poor performance with data from larger cities only. Note that the negative coefficient indicates that there are fewer poor performers in larger cities.

hit rate threshold	β	low	high
0.450	-0.037	-0.047	-0.026
0.460	-0.036	-0.047	-0.025
0.470	-0.036	-0.047	-0.026
0.480	-0.036	-0.046	-0.025
0.490	-0.036	-0.046	-0.025
0.500	-0.036	-0.046	-0.025
0.510	-0.010	-0.017	-0.004
0.520	-0.010	-0.016	-0.005
0.530	-0.010	-0.016	-0.004
0.540	-0.010	-0.016	-0.004
0.550	-0.009	-0.015	-0.004

Supplementary Table 7.8: Sensitivity test for the threshold for “perfect" performance with all data included.

hit rate threshold	β	low	high
0.850	0.088	0.072	0.113
0.860	0.088	0.070	0.112
0.870	0.088	0.072	0.112
0.880	0.097	0.077	0.132
0.890	0.093	0.072	0.141
0.900	0.093	0.070	0.134
0.910	0.093	0.068	0.135
0.920	0.103	0.078	0.169
0.930	0.103	0.075	0.162
0.940	0.103	0.075	0.160
0.950	0.100	0.074	0.156

Supplementary Table 7.9: Sensitivity test for the threshold for “perfect" performance with data from larger cities only.

hit rate threshold	β	low	high
0.850	0.109	0.085	0.135
0.860	0.109	0.084	0.134
0.870	0.109	0.086	0.132
0.880	0.142	0.115	0.173
0.890	0.163	0.129	0.202
0.900	0.163	0.132	0.197
0.910	0.163	0.129	0.199
0.920	0.170	0.134	0.217
0.930	0.170	0.131	0.215
0.940	0.170	0.131	0.216
0.950	0.171	0.131	0.214

REFERENCES

Substance Abuse and Mental Health Services Administration. Substance abuse and mental health services administration, results from the 2010 national survey on drug use and health: Summary of national findings. *NSDUH Series H-41, HHS Publication*, 1((SMA): 11–4658, 2011.

Ahmad Alabdulkareem, Morgan R Frank, Lijun Sun, Bedoor AlShebli, César Hidalgo, and Iyad Rahwan. Unpacking the polarization of workplace skills. *Science advances*, 4(7): eaao6030, 2018.

Gordon Willard Allport, Kenneth Clark, and Thomas Pettigrew. *The nature of prejudice*. Addison-wesley Reading, MA, 1954.

Albandri Alotaibi, Geoffrey Underwood, and Alastair D Smith. Cultural differences in attention: Eye movement evidence from a comparative visual search task. *Consciousness and cognition*, 55:254–265, 2017.

Luiz GA Alves, Haroldo V Ribeiro, Ervin K Lenzi, and Renio S Mendes. Distance to the scaling law: a useful approach for unveiling relationships between crime and urban metrics. *PloS one*, 8(8):e69580, 2013.

David M Amodio. The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience*, 15(10):670–682, 2014.

Reinaldo B Arellano-Valle and Adelchi Azzalini. The centred parameterization and related quantities of the skew-t distribution. *Journal of Multivariate Analysis*, 113:73–90, 2013.

Rudy Arthur and Hywel TP Williams. Scaling laws in geo-located twitter data. *PloS one*, 14(7):e0218454, 2019.

Joseph B Bak-Coleman, Mark Alfano, Wolfram Barfuss, Carl T Bergstrom, Miguel A Centeno, Iain D Couzin, Jonathan F Donges, Mirta Galesic, Andrew S Gersick, Jennifer Jacquet, et al. Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences*, 118(27):e2025764118, 2021.

Albert-László Barabási et al. *Network science*. Cambridge university press, 2016.

Daniel Barkoczi and Mirta Galesic. Social learning strategies modify the effect of network structure on group performance. *Nature communications*, 7(1):1–8, 2016.

Andrew Scott Baron and Mahzarin R Banaji. The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological science*, 17(1): 53–58, 2006.

Michael S Barton, Frederick Weil, Melinda Jackson, and Darien A Hickey. An investigation of the influence of the spatial distribution of neighborhood violent crime on fear of crime. *Crime & Delinquency*, 63(13):1757–1776, 2017.

Douglas Bates, Deepayan Sarkar, Maintainer Douglas Bates, and L Matrix. The lme4 package. *R package version*, 2(1):74, 2007.

Marc G Berman, John Jonides, and Stephen Kaplan. The cognitive benefits of interacting with nature. *Psychological science*, 19(12):1207–1212, 2008.

Marc G Berman, Andrew J Stier, and Gaby N Akcelik. Environmental neuroscience. *American Psychologist*, 74(9):1039, 2019.

Luis Bettencourt and Geoffrey West. A unified theory of urban living. *Nature*, 467(7318): 912–913, 2010.

Luis M. A. Bettencourt. *Introduction to urban science: evidence and theory of cities as complex systems*. The MIT Press, Cambridge, Massachusetts, 2021a. ISBN 978-0-262-04600-8.

Luís MA Bettencourt. The origins of scaling in cities. *Science*, 340(6139):1438–1441, 2013.

Luis MA Bettencourt. Urban growth and the emergent statistics of cities. *Science advances*, 6(34):eaat8812, 2020.

Luís MA Bettencourt. *Introduction to urban science: evidence and theory of cities as complex systems*. MIT Press, 2021b.

Luís MA Bettencourt, José Lobo, Dirk Helbing, Christian Kühnert, and Geoffrey B West. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the national academy of sciences*, 104(17):7301–7306, 2007a.

Luis MA Bettencourt, Jose Lobo, and Deborah Strumsky. Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size. *Research policy*, 36(1): 107–120, 2007b.

Luís MA Bettencourt, José Lobo, Deborah Strumsky, and Geoffrey B West. Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PloS one*, 5(11):e13541, 2010.

Luís MA Bettencourt, Vicky Chuqiao Yang, José Lobo, Christopher P Kempes, Diego Rybski, and Marcus J Hamilton. The interpretation of urban scaling analysis in time. *Journal of the Royal Society Interface*, 17(163):20190846, 2020.

Carlos Blanco, Mayumi Okuda, John C Markowitz, Shang-Min Liu, Bridget F Grant, and Deborah S Hasin. The epidemiology of chronic major depressive disorder and dysthymic disorder: results from the national epidemiologic survey on alcohol and related conditions. *The Journal of clinical psychiatry*, 71(12):1645, 2010.

Giovanni Bonaccorsi, Francesco Pierri, Matteo Cinelli, Andrea Flori, Alessandro Galeazzi, Francesco Porcelli, Ana Lucia Schmidt, Carlo Michele Valensise, Antonio Scala, Walter Quattrociocchi, and Fabio Pammolli. Economic and social consequences of human mobility restrictions under COVID-19. *PNAS*, 117(27):1553015535, July 2020.

Christy K Boscardin. Reducing implicit bias through curricular interventions. *Journal of General Internal Medicine*, 30(12):1726–1728, 2015.

Vincent Boucher. Structural homophily. *International Economic Review*, 56(1):235–264, 2015.

Tobias Brosch, Eyal Bar-David, and Elizabeth A Phelps. Implicit race bias decreases the similarity of neural representations of black and white faces. *Psychological science*, 24(2): 160–166, 2013.

Alexander P Burgoyne and Randall W Engle. Attention control: A cornerstone of higher-order cognition. *Current Directions in Psychological Science*, 29(6):624–630, 2020.

Richard A. Burns, Kaarin J. Anstey, and Timothy D. Windsor. Subjective Well-Being Mediates the Effects of Resilience and Mastery on Depression and Anxiety in a Large Community Sample of Young and Middle-Aged Adults. *Australian & New Zealand Journal of Psychiatry*, 45(3):240–248, March 2011. ISSN 0004-8674, 1440-1614. doi:10.3109/00048674.2010.529604. URL `http://journals.sagepub.com/doi/10.3109/00048674.2010.529604`.

Jedelyn Cabrieto, Francis Tuerlinckx, Peter Kuppens, Mariel Grassmann, and Eva Ceulemans. Detecting correlation changes in multivariate time series: A comparison of four non-parametric change point detection methods. *Behavior Research Methods*, 49(3):988–1005, 2017.

John T Cacioppo and Louise C Hawkley. Perceived social isolation and cognition. *Trends in cognitive sciences*, 13(10):447–454, 2009.

John T Cacioppo, James H Fowler, and Nicholas A Christakis. Alone in the crowd: the structure and spread of loneliness in a large social network. *Journal of personality and social psychology*, 97(6):977, 2009.

Serge Caparos, Karina J Linnell, Andrew J Bremner, Jan W de Fockert, and Jules Davidoff. Do local and global perceptual biases tell us anything about local and global selective attention? *Psychological science*, 24(2):206–212, 2013.

Rudolf Cesaretti, José Lobo, Luís MA Bettencourt, Scott G Ortman, and Michael E Smith. Population-area relationship for medieval european cities. *PloS one*, 11(10):e0162678, 2016.

David H Chae, Amani M Nuru-Jeter, and Nancy E Adler. Implicit racial bias as a moderator of the association between racial discrimination and hypertension: a study of midlife african american men. *Psychosomatic Medicine*, 74(9):961, 2012.

David H Chae, Wizdom A Powell, Amani M Nuru-Jeter, Mia A Smith-Bynum, Eleanor K Seaton, Tyrone A Forman, Rodman Turpin, and Robert Sellers. The role of racial identity and implicit racial bias in self-reported racial discrimination: Implications for depression among african american men. *Journal of Black Psychology*, 43(8):789–812, 2017.

Taylor A Chamberlain and Monica D Rosenberg. Propofol selectively modulates functional connectivity signatures of sustained attention during rest and narrative listening. *Cerebral Cortex*, 32(23):5362–5375, 2022.

Juan Chen, Deborah S. Davis, Kaming Wu, and Haijing Dai. Life satisfaction in urbanizing China: The effect of city size and pathways to urban residency. *Cities*, 49:88–97, December 2015. ISSN 02642751. doi:10.1016/j.cities.2015.07.011. URL `https://linkinghub.elsevier.com/retrieve/pii/S0264275115001110`.

Raj Chetty, Matthew O Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, et al. Social capital i: measurement and associations with economic mobility. *Nature*, 608(7921): 108–121, 2022a.

Raj Chetty, Matthew O Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, et al. Social capital ii: determinants of economic connectedness. *Nature*, 608(7921):122–134, 2022b.

Andy Clark. *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press, 2015.

Luke Clark and Guy M Goodwin. State-and trait-related deficits in sustained attention in bipolar disorder. *European archives of psychiatry and clinical neuroscience*, 254:61–68, 2004.

Jasmin Cloutier, Tianyi Li, Bratislav Mišić, Joshua Correll, and Marc G Berman. Brain network activity during face perception: the impact of perceptual familiarity and individual differences in childhood experience. *Cerebral Cortex*, 27(9):4326–4338, 2017.

Antoine Coutrot, Ed Manley, Sarah Goodroe, Christoffer Gahnstrom, Gabriele Filomena, Demet Yesiltepe, RC Dalton, Jan M Wiener, Christian Hölscher, Michael Hornberger, et al. Entropy of city street networks linked to future spatial navigation ability. *Nature*, 604(7904):104–110, 2022.

Edward RFW Crossman. A theory of the acquisition of speed-skill. *Ergonomics*, 2(2): 153–166, 1959.

Robert A. Cummins. Subjective Wellbeing, Homeostatically Protected Mood and Depression: A Synthesis. *Journal of Happiness Studies*, 11(1):1–17, March 2010. ISSN 1389-4978, 1573-7780. doi:10.1007/s10902-009-9167-0. URL `http://link.springer.com/10.1007/s10902-009-9167-0`.

Raphael E Cuomo, Vidya Purushothaman, Jiawei Li, Mingxiang Cai, and Tim K Mackey. A longitudinal and geospatial analysis of covid-19 tweets during the early outbreak period in the united states. *BMC Public Health*, 21(1):1–11, 2021.

Anders M. Dale, Bruce Fischl, and Martin I. Sereno. Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, 9(2):179–194, 1999. ISSN 1053-8119. doi:10.1006/nimg.1998.0395. URL http://www.sciencedirect.com/science/article/pii/S1053811998903950.

Jonas Dalege and Tamara van der Does. Using a cognitive network model of moral and social beliefs to explain belief change. *Science Advances*, 8(33):eabm0137, 2022. doi:10.1126/sciadv.abm0137. URL https://www.science.org/doi/abs/10.1126/sciadv.abm0137.

Jonas Dalege, Denny Borsboom, Frenk van Harreveld, Lourens J Waldorp, and Han LJ van der Maas. Network structure explains the impact of attitudes on voting decisions. *Scientific reports*, 7(1):1–11, 2017.

Luca Dall'Asta, Claudio Castellano, and Matteo Marsili. Statistical physics of the schelling model of segregation. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(07): L07002, 2008.

Yunxiao Dang, Li Chen, Wenzhong Zhang, Dan Zheng, and Dongsheng Zhan. How does growing city size affect residents' happiness in urban China? A case study of the Bohai rim area. *Habitat International*, 97:102120, March 2020. ISSN 01973975. doi:10.1016/j.habitatint.2020.102120. URL https://linkinghub.elsevier.com/retrieve/pii/S019739751930267X.

Geraldine Dawson, Karen Toth, Robert Abbott, Julie Osterling, Jeff Munson, Annette Estes, and Jane Liaw. Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental psychology*, 40(2):271, 2004.

Jan De Houwer. Implicit bias is behavior: A functional-cognitive perspective on implicit bias. *Perspectives on Psychological Science*, 14(5):835–840, 2019.

Morteza Dehghani, Kate Johnson, Joe Hoover, Eyal Sagi, Justin Garten, Niki Jitendra Parmar, Stephen Vaisey, Rumen Iliev, and Jesse Graham. Purity homophily in social networks. *Journal of Experimental Psychology: General*, 145(3):366, 2016.

Erin Dehon, Nicole Weiss, Jonathan Jones, Whitney Faulconer, Elizabeth Hinton, and Sarah Sterling. A systematic review of the impact of physician implicit racial bias on clinical decision making. *Academic Emergency Medicine*, 24(8):895–904, 2017.

Emre Demiralp, Renee J. Thompson, Jutta Mata, Susanne M. Jaeggi, Martin Buschkuehl, Lisa Feldman Barrett, Phoebe C. Ellsworth, Metin Demiralp, Luis Hernandez-Garcia, Patricia J. Deldin, Ian H. Gotlib, and John Jonides. Feeling Blue or Turquoise? Emotional Differentiation in Major Depressive Disorder. *Psychological Science*, 23(11):1410–1416, November 2012. ISSN 0956-7976, 1467-9280. doi:10.1177/0956797612444903. URL http://journals.sagepub.com/doi/10.1177/0956797612444903.

Ed Diener and Alex C. Michalos, editors. *The Science of Well-Being*, volume 37 of *Social Indicators Research Series*. Springer Netherlands, Dordrecht, 2009. ISBN 978-90-481-2349-0 978-90-481-2350-6. doi:10.1007/978-90-481-2350-6. URL `http://link.springer.com/10.1007/978-90-481-2350-6`.

Brian M. D'Onofrio, Arvid Sjölander, Benjamin B. Lahey, Paul Lichtenstein, and A. Sara Öberg. Accounting for Confounding in Observational Studies. *Annual Review of Clinical Psychology*, 16(1):25–48, May 2020. ISSN 1548-5943, 1548-5951. doi:10.1146/annurev-clinpsy-032816-045030. URL `https://www.annualreviews.org/doi/10.1146/annurev-clinpsy-032816-045030`.

Yarrow Dunham, Andrew Scott Baron, and Mahzarin R Banaji. From american city to japanese village: A cross-cultural investigation of implicit race attitudes. *Child development*, 77(5):1268–1281, 2006.

Jacob Eisenstein, Brendan O'Connor, Noah A Smith, and Eric Xing. A latent variable model for geographic lexical variation. In *Proceedings of the 2010 conference on empirical methods in natural language processing*, pages 1277–1287, 2010.

Pierce D Ekstrom, Joel M Le Forestier, and Calvin K Lai. Racial demographics explain the link between racial disparities in traffic stops and county-level racial attitudes. *Psychological science*, 33(4):497–509, 2022.

Ingrid Gould Ellen and Katherine O'Regan. Crime and urban flight revisited: The effect of the 1990s drop in crime on cities. *Journal of Urban Economics*, 68(3):247–259, 2010.

Gokhan Ertug, Julia Brennecke, Balázs Kovács, and Tengjian Zou. What does homophily do? a review of the consequences of homophily. *Academy of Management Annals*, 16(1): 38–69, 2022.

Oscar Esteban, Ross Blair, Christopher J. Markiewicz, Shoshana L. Berleant, Craig Moodie, Feilong Ma, Ayse Ilkay Isik, Asier Erramuzpe, Mathias Kent, James D. andGoncalves, Elizabeth DuPre, Kevin R. Sitek, Daniel E. P. Gomez, Daniel J. Lurie, Zhifang Ye, Russell A. Poldrack, and Krzysztof J. Gorgolewski. fmriprep 22.1.0+0.gce344b39.dirty. *Software*, 2018a. doi:10.5281/zenodo.852659.

Oscar Esteban, Christopher Markiewicz, Ross W Blair, Craig Moodie, Ayse Ilkay Isik, Asier Erramuzpe Aliaga, James Kent, Mathias Goncalves, Elizabeth DuPre, Madeleine Snyder, Hiroyuki Oya, Satrajit Ghosh, Jessey Wright, Joke Durnez, Russell Poldrack, and Krzysztof Jacek Gorgolewski. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, 2018b. doi:10.1038/s41592-018-0235-4.

Oscar Esteban, Christopher J Markiewicz, Ross W Blair, Craig A Moodie, A Ilkay Isik, Asier Erramuzpe, James D Kent, Mathias Goncalves, Elizabeth DuPre, Madeleine Snyder, et al. fmriprep: a robust preprocessing pipeline for functional mri. *Nature methods*, 16(1):111–116, 2019.

Stephen Eubank, Hasan Guclu, VS Anil Kumar, Madhav V Marathe, Aravind Srinivasan, Zoltan Toroczkai, and Nan Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429(6988):180–184, 2004.

Thomas A. Fergus and Joseph R. Bardeen. Negative mood regulation expectancies moderate the association between happiness emotion goals and depressive symptoms. *Personality and Individual Differences*, 100:23–27, October 2016. ISSN 01918869. doi:10.1016/j.paid.2015.08.010. URL `https://linkinghub.elsevier.com/retrieve/pii/S0191886915005152`.

Anna V Fisher. Selective sustained attention: A developmental foundation for cognition. *Current opinion in psychology*, 29:248–253, 2019.

Brett Q. Ford, Amanda J. Shallcross, Iris B. Mauss, Victoria A. Floerke, and June Gruber. Desperately Seeking Happiness: Valuing Happiness is Associated With Symptoms and Diagnosis of Depression. *Journal of Social and Clinical Psychology*, 33(10):890–905, December 2014. ISSN 0736-7236. doi:10.1521/jscp.2014.33.10.890. URL `http://guilfordjournals.com/doi/10.1521/jscp.2014.33.10.890`.

Francesca C Fortenbaugh, Joseph DeGutis, Laura Germine, Jeremy B Wilmer, Mallory Grosso, Kathryn Russo, and Michael Esterman. Sustained attention across the life span in a sample of 10,000: Dissociating ability and strategy. *Psychological science*, 26(9): 1497–1510, 2015.

Mirta Galesic, Henrik Olsson, Jonas Dalege, Tamara van der Does, and Daniel L Stein. Integrating social and cognitive aspects of belief dynamics: towards a unifying framework. *Journal of the Royal Society Interface*, 18(176):20200857, 2021.

Bentley L Gibson, Philippe Rochat, Erin B Tone, and Andrew S Baron. Sources of implicit and explicit intergroup race bias among african-american children and young adults. *PloS one*, 12(9):e0183015, 2017.

Stephen E Gilman, Ichiro Kawachi, Garrett M Fitzmaurice, and Stephen L Buka. Socio-economic status, family disruption and residential stability in childhood: relation to onset, recurrence and remission of major depression. *Psychological medicine*, 33(8):1341, 2003.

Edward L. Glaeser, Joshua D. Gottlieb, and Oren Ziv. Unhappy Cities. *Journal of Labor Economics*, 34(S2):S129–S182, April 2016. ISSN 0734-306X, 1537-5307. doi:10.1086/684044. URL `https://www.journals.uchicago.edu/doi/10.1086/684044`.

Benjamin Golub and Matthew O Jackson. How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics*, 127(3):1287–1338, 2012.

Antonya M Gonzalez, Jennifer R Steele, and Andrew S Baron. Reducing children's implicit racial bias through exposure to positive out-group exemplars. *Child Development*, 88(1): 123–130, 2017.

K. Gorgolewski, C. D. Burns, C. Madison, D. Clark, Y. O. Halchenko, M. L. Waskom, and S. Ghosh. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, 5:13, 2011. doi:10.3389/fninf.2011.00013.

Krzysztof J. Gorgolewski, Oscar Esteban, Christopher J. Markiewicz, Erik Ziegler, David Gage Ellis, Michael Philipp Notter, Dorota Jarecka, Hans Johnson, Christopher Burns, Alexandre Manhães-Savio, Carlo Hamalainen, Benjamin Yvernault, Taylor Salo, Kesshi Jordan, Mathias Goncalves, Michael Waskom, Daniel Clark, Jason Wong, Fred Loney, Marc Modat, Blake E Dewey, Cindee Madison, Matteo Visconti di Oleggio Castello, Michael G. Clark, Michael Dayan, Dav Clark, Anisha Keshavan, Basile Pinsard, Alexandre Gramfort, Shoshana Berleant, Dylan M. Nielson, Salma Bougacha, Gael Varoquaux, Ben Cipollini, Ross Markello, Ariel Rokem, Brendan Moloney, Yaroslav O. Halchenko, Demian Wassermann, Michael Hanke, Christian Horea, Jakub Kaczmarzyk, Gilles de Hollander, Elizabeth DuPre, Ashley Gillman, David Mordom, Colin Buchanan, Rosalia Tungaraza, Wolfgang M. Pauli, Shariq Iqbal, Sharad Sikka, Matteo Mancini, Yannick Schwartz, Ian B. Malone, Mathieu Dubois, Caroline Frohlich, David Welch, Jessica Forbes, James Kent, Aimi Watanabe, Chad Cumba, Julia M. Huntenburg, Erik Kastman, B. Nolan Nichols, Arman Eshaghi, Daniel Ginsburg, Alexander Schaefer, Benjamin Acland, Steven Giavasis, Jens Kleesiek, Drew Erickson, René Küttner, Christian Haselgrove, Carlos Correa, Ali Ghayoor, Franz Liem, Jarrod Millman, Daniel Haehn, Jeff Lai, Dale Zhou, Ross Blair, Tristan Glatard, Mandy Renfro, Siqi Liu, Ari E. Kahn, Fernando Pérez-García, William Triplett, Leonie Lampe, Jörg Stadler, Xiang-Zhen Kong, Michael Hallquist, Andrey Chetverikov, John Salvatore, Anne Park, Russell Poldrack, R. Cameron Craddock, Souheil Inati, Oliver Hinds, Gavin Cooper, L. Nathan Perkins, Ana Marina, Aaron Mattfeld, Maxime Noel, Lukas Snoek, K Matsubara, Brian Cheung, Simon Rothmei, Sebastian Urchs, Joke Durnez, Fred Mertz, Daniel Geisler, Andrew Floren, Stephan Gerhard, Paul Sharp, Miguel Molina-Romero, Alejandro Weinstein, William Broderick, Victor Saase, Sami Kristian Andberg, Robbert Harms, Kai Schlamp, Jaime Arias, Dimitri Papadopoulos Orfanos, Claire Tarbert, Arielle Tambini, Alejandro De La Vega, Thomas Nickson, Matthew Brett, Marcel Falkiewicz, Kornelius Podranski, Janosch Linkersdörfer, Guillaume Flandin, Eduard Ort, Dmitry Shachnev, Daniel McNamee, Andrew Davison, Jan Varada, Isaac Schwabacher, John Pellman, Martin Perez-Guevara, Ranjeet Khanuja, Nicolas Pannetier, Conor McDermottroe, and Satrajit Ghosh. Nipype. *Software*, 2018. doi:10.5281/zenodo.596855.

Jack M. Gorman. Comorbid depression and anxiety spectrum disorders. *Depression and Anxiety*, 4(4):160–168, 1996. doi:10.1002/(SICI)1520-6394(1996)4:4<160::AID-DA2>3.0.CO;2-J.

Corina Graif, Alina Lungeanu, and Alyssa M Yetter. Neighborhood isolation in chicago: Violent crime effects on structural isolation and homophily in inter-neighborhood commuting networks. *Social networks*, 51:40–59, 2017a.

Corina Graif, Alina Lungeanu, and Alyssa M. Yetter. Neighborhood isolation in Chicago:

Violent crime effects on structural isolation and homophily in inter-neighborhood commuting networks. *Social Networks*, 51:40–59, October 2017b. ISSN 03788733. doi:10.1016/j.socnet.2017.01.007. URL https://linkinghub.elsevier.com/retrieve/pii/S0378873316302179.

Mark S Granovetter. The strength of weak ties. *American journal of sociology*, 78(6): 1360–1380, 1973.

C Shawn Green and Daphne Bavelier. Learning, attentional control, and action video games. *Current biology*, 22(6):R197–R206, 2012.

Paul E Greenberg, Ronald C Kessler, Howard G Birnbaum, Stephanie A Leong, Sarah W Lowe, Patricia A Berglund, and Patricia K Corey-Lisle. The economic burden of depression in the united states: how did it change between 1990 and 2000? *Journal of clinical psychiatry*, 64(12):1465–1475, 2003.

Ross W Greene, Joseph Biederman, Stephen V Faraone, Michael C Monuteaux, Eric Mick, EMILY P DuPRE, Catherine S Fine, and Jennifer C Goring. Social impairment in girls with adhd: patterns, gender comparisons, and correlates. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(6):704–710, 2001.

Anthony G Greenwald, Brian A Nosek, and Mahzarin R Banaji. Understanding and using the implicit association test: I. an improved scoring algorithm. *Journal of personality and social psychology*, 85(2):197, 2003.

Anthony G Greenwald, Mahzarin R Banaji, and Brian A Nosek. Statistically small effects of the implicit association test can have societally large effects. *Journal of Personality and Social Pyschology*, 2015.

Oliver Gruebner, Michael A Rapp, Mazda Adli, Ulrike Kluge, Sandro Galea, and Andreas Heinz. Cities and mental health. *Deutsches Ärzteblatt International*, 114(8):121, 2017.

Bernard Guyer, Mary Anne Freedman, Donna M Strobino, and Edward J Sondik. Annual summary of vital statistics: trends in the health of americans during the 20th century. *Pediatrics*, 106(6):1307–1317, 2000.

Yosh Halberstam and Brian Knight. Homophily, group size, and the diffusion of political information in social networks: Evidence from twitter. *Journal of public economics*, 143: 73–88, 2016.

Thomas Hansen and Britt Slagsvold. The age and subjective well-being paradox revisited:A multidimensional perspective. *Norsk Epidemiologi*, 22(2), November 2012. ISSN 0803-2491. doi:10.5324/nje.v22i2.1565. URL http://www.ntnu.no/ojs/index.php/norepid/article/view/1565.

Sarra Hedden, Joe Gfroerer, Peggy Barker, Shelagh Smith, Michael R Pemberton, Lissette M Saavedra, Valerie L Forman-Hoffman, Heather Ringeisen, and Scott P Novak. Comparison

of nsduh mental health data and methods with other data sources. In *CBHSQ Data Review*. Substance Abuse and Mental Health Services Administration (US), 2012.

Eric Hehman, Jessica K Flake, and Jimmy Calanchini. Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *Social psychological and personality science*, 9(4):393–401, 2018.

Eric Hehman, Eugene K Ofosu, and Jimmy Calanchini. Using environmental features to maximize prediction of regional intergroup bias. *Social Psychological and Personality Science*, 12(2):156–164, 2021.

Cate Heine, Cristina Marquez, Paolo Santi, Marcus Sundberg, Miriam Nordfors, and Carlo Ratti. Analysis of mobility homophily in stockholm based on social network data. *PloS one*, 16(3):e0247996, 2021.

Joseph Henrich. *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. princeton University press, 2016.

Marek Hlavac. Stargazer: Well-formatted regression and summary statistics tables. *R package version*, 5(1), 2015.

Liangjie Hong and Brian D Davison. Empirical study of topic modeling in twitter. In *Proceedings of the first workshop on social media analytics*, pages 80–88, 2010.

Roxane S Hoyer, Eric Pakulak, Aurélie Bidet-Caulet, and Christina M Karns. Differences in sustained attention but not distraction in preschoolers from lower socioeconomic status backgrounds. *BioRxiv*, pages 2021–04, 2021.

Betsy Hoza, Sylvie Mrug, Alyson C Gerdes, Stephen P Hinshaw, William M Bukowski, Joel A Gold, Helena C Kraemer, William E Pelham Jr, Timothy Wigal, and L Eugene Arnold. What aspects of peer relationships are impaired in children with attention-deficit/hyperactivity disorder? *Journal of consulting and clinical psychology*, 73(3):411, 2005.

Liqiang Huang, Lei Mo, and Ying Li. Measuring the interrelations among multiple paradigms of visual attention: an individual differences approach. *Journal of experimental psychology: human perception and performance*, 38(2):414, 2012.

Mia Hubert and Michiel Debruyne. Minimum covariance determinant. *Wiley interdisciplinary reviews: Computational statistics*, 2(1):36–43, 2010.

Baqar A Husaini, Stephen T Moore, and Robert S Castor. Social and psychological well-being of black elderly living in high-rises for the elderly. *Journal of Gerontological Social Work*, 16(3-4):57–78, 1991.

Ryan Hyon, Yoosik Youm, Junsol Kim, Jeanyung Chey, Seyul Kwak, and Carolyn Parkinson. Similarity in functional brain connectivity at rest predicts interpersonal closeness in the

social network of an entire village. *Proceedings of the National Academy of Sciences*, 117 (52):33149–33160, 2020.

Stefano Maria Iacus, Giuseppe Porro, Silvia Salini, and Elena Siletti. Social networks data and subjective well-being. an innovative measurement for italian provinces. *Scienze Regionali*, 18(Speciale):667–678, 2019.

Herminia Ibarra. Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative science quarterly*, pages 422–447, 1992.

Yoshio Itaba. Does City Size Affect Happiness? In Toshiaki Tachibanaki, editor, *Advances in Happiness Research: A Comparative Perspective*, pages 245–273. Springer Japan, Tokyo, 2016. ISBN 978-4-431-55753-1. doi:10.1007/978-4-431-55753-1_14. URL `https://doi.org/10.1007/978-4-431-55753-1_14`.

Matthew O Jackson. Inequality's economic and social roots: The role of social networks and homophily. *Available at SSRN 3795626*, 2021.

Sarah M Jackson, Amy L Hillard, and Tamera R Schneider. Using implicit bias training to improve attitudes toward women in stem. *Social Psychology of Education*, 17(3):419–438, 2014.

Jane Jacobs. *The death and life of great American cities*. Vintage, 2016.

Drew S Jacoby-Senghor, Stacey Sinclair, and J Nicole Shelton. A lesson in bias: The relationship between implicit racial bias and performance in pedagogical contexts. *Journal of Experimental Social Psychology*, 63:50–55, 2016.

Spencer L James, Degu Abate, Kalkidan Hassen Abate, Solomon M Abay, Cristiana Abbafati, Nooshin Abbasi, Hedayat Abbastabar, Foad Abd-Allah, Jemal Abdela, Ahmed Abdelalim, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, 392(10159): 1789–1858, 2018.

Łukasz D. Kaczmarek, Błażej Bączkowski, Jolanta Enko, Barbara Baran, and Peter Theuns. Subjective well-being as a mediator for curiosity and depression. *Polish Psychological Bulletin*, 45(2):200–204, June 2014. ISSN 1641-7844. doi:10.2478/ppb-2014-0025. URL `http://journals.pan.pl/dlibra/publication/114748/edition/99801/content`.

Julia B Kahn, Ryan D Ward, Lora W Kahn, Nicole M Rudy, Eric R Kandel, Peter D Balsam, and Eleanor H Simpson. Medial prefrontal lesions in mice impair sustained attention but spare maintenance of information in working memory. *Learning & Memory*, 19(11):513–517, 2012.

Stephen Kaplan. The restorative benefits of nature: Toward an integrative framework. *Journal of environmental psychology*, 15(3):169–182, 1995.

Omid Kardan, Laura Shneidman, Sheila Krogh-Jespersen, Suzanne Gaskins, Marc G Berman, and Amanda Woodward. Cultural and developmental influences on overt visual attention to videos. *Scientific reports*, 7(1):1–16, 2017.

Omid Kardan, Andrew J Stier, Carlos Cardenas-Iniguez, Kathryn E Schertz, Julia C Pruin, Yuting Deng, Taylor Chamberlain, Wesley J Meredith, Xihan Zhang, Jillian E Bowman, et al. Differences in the functional brain architecture of sustained attention and working memory in youth and adults. *Plos Biology*, 20(12):e3001938, 2022.

Fariba Karimi, Mathieu Génois, Claudia Wagner, Philipp Singer, and Markus Strohmaier. Homophily influences ranking of minorities in social networks. *Scientific reports*, 8(1): 1–12, 2018.

Andrea Kavanaugh, Debbie Denise Reese, John M Carroll, and Mary Beth Rosson. Weak ties in networked communities. In *Communities and Technologies: Proceedings of the First International Conference on Communities and Technologies; C&T 2003*, pages 265–286. Springer, 2003.

Ade Kearns, Elise Whitley, Phil Mason, and Lyndal Bond. 'living the high life'? residential, social and psychosocial outcomes for high-rise occupants in a deprived context. *Housing Studies*, 27(1):97–126, 2012.

Mohammedhamid Osman Kelifa, Yinmei Yang, Herbert Carly, Wang Bo, and Peigang Wang. How Adverse Childhood Experiences Relate to Subjective Wellbeing in College Students: The Role of Resilience and Depression. *Journal of Happiness Studies*, September 2020. ISSN 1389-4978, 1573-7780. doi:10.1007/s10902-020-00308-7. URL `http://link.springer.com/10.1007/s10902-020-00308-7`.

Jun Won Kim, Bung-Nyun Kim, Johanna Inhyang Kim, Young Sik Lee, Kyung Joon Min, Hyun-Jin Kim, and Jaewon Lee. Social network analysis reveals the negative effects of attention-deficit/hyperactivity disorder (adhd) symptoms on friend-based student networks. *PloS one*, 10(11):e0142782, 2015.

Amy JH Kind and William R Buckingham. Making neighborhood-disadvantage metrics accessible—the neighborhood atlas. *The New England journal of medicine*, 378(26):2456, 2018.

Arno Klein, Satrajit S. Ghosh, Forrest S. Bao, Joachim Giard, Yrjö Häme, Eliezer Stavsky, Noah Lee, Brian Rossa, Martin Reuter, Elias Chaibub Neto, and Anisha Keshavan. Mindboggling morphometry of human brains. *PLOS Computational Biology*, 13(2):e1005350, 2017. ISSN 1553-7358. doi:10.1371/journal.pcbi.1005350. URL `http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005350`.

Jeffrey T Klein, Stephen V Shepherd, and Michael L Platt. Social attention and the brain. *Current Biology*, 19(20):R958–R962, 2009.

Heli Koivumaa-Honkanen, Jaakko Kaprio, Risto Honkanen, Heimo Viinamki, and Markku Koskenvuo. Life satisfaction and depression in a 15-year follow-up of healthy adults. *Social Psychiatry and Psychiatric Epidemiology*, 39(12):994–999, December 2004. ISSN 0933-7954, 1433-9285. doi:10.1007/s00127-004-0833-6. URL `http://link.springer.com/10.1007/s00127-004-0833-6`.

Gueorgi Kossinets and Duncan J Watts. Origins of homophily in an evolving social network. *American journal of sociology*, 115(2):405–450, 2009.

Moritz Köster, Shoji Itakura, Relindis Yovsi, and Joscha Kärtner. Visual attention in 5-year-olds from three different cultures. *PloS one*, 13(7):e0200239, 2018.

Lydia Krabbendam and Jim Van Os. Schizophrenia and urbanicity: a major environmental influence—conditional on genetic risk. *Schizophrenia bulletin*, 31(4):795–799, 2005.

Michael R Kramer and Carol R Hogue. Is segregation bad for your health? *Epidemiologic reviews*, 31(1):178–194, 2009.

Kevin M Kruse. White flight. In *White Flight*. Princeton University Press, 2013.

Jennifer T Kubota, Mahzarin R Banaji, and Elizabeth A Phelps. The neuroscience of race. *Nature neuroscience*, 15(7):940–948, 2012.

Agustin Lage-Castellanos, Giancarlo Valente, Elia Formisano, and Federico De Martino. Methods for computing the maximum performance of computational models of fmri responses. *PLoS computational biology*, 15(3):e1006397, 2019.

Calvin K Lai, Maddalena Marini, Steven A Lehr, Carlo Cerruti, Jiyun-Elizabeth L Shin, Jennifer A Joy-Gaba, Arnold K Ho, Bethany A Teachman, Sean P Wojcik, Spassena P Koleva, et al. Reducing implicit racial preferences: I. a comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143(4):1765, 2014.

Calvin K Lai, Allison L Skinner, Erin Cooley, Sohad Murrar, Markus Brauer, Thierry Devos, Jimmy Calanchini, Y Jenny Xiao, Christina Pedram, Christopher K Marshburn, et al. Reducing implicit racial preferences: Ii. intervention effectiveness across time. *Journal of Experimental Psychology: General*, 145(8):1001, 2016.

Neal Lathia, Daniele Quercia, and Jon Crowcroft. The hidden image of the city: sensing community well-being from urban mobility. In *International conference on pervasive computing*, pages 91–98. Springer, 2012a.

Neal Lathia, Daniele Quercia, and Jon Crowcroft. The Hidden Image of the City: Sensing Community Well-Being from Urban Mobility. In Judy Kay, Paul Lukowicz, Hideyuki Tokuda, Patrick Olivier, and Antonio Krüger, editors, *Pervasive Computing*, pages 91–98, Berlin, Heidelberg, 2012b. Springer Berlin Heidelberg. ISBN 978-3-642-31205-2.

James Laurence. The effect of ethnic diversity and community disadvantage on social cohesion: A multi-level analysis of social capital and interethnic relations in uk communities. *European Sociological Review*, 27(1):70–89, 2011.

Florian Lederbogen, Peter Kirsch, Leila Haddad, Fabian Streit, Heike Tost, Philipp Schuch, Stefan Wüst, Jens C Pruessner, Marcella Rietschel, Michael Deuschle, et al. City living and urban upbringing affect neural social stress processing in humans. *Nature*, 474(7352): 498–501, 2011.

Eun Lee, Fariba Karimi, Claudia Wagner, Hang-Hyun Jo, Markus Strohmaier, and Mirta Galesic. Homophily and minority-group size explain perception biases in social networks. *Nature human behaviour*, 3(10):1078–1087, 2019.

Hyunhee Lee. Psychological characteristics of high-rise residents. *International Journal of Sustainable Building Technology and Urban Development*, 5(1):10–20, 2014.

Ju-Mi Lee, Won Joon Lee, Hyeon Chang Kim, Wungrak Choi, Jina Lee, Kiho Sung, Sang Hui Chu, Yeong-Ran Park, and Yoosik Youm. The korean social life, health and aging project-health examination cohort. *Epidemiology and health*, 36, 2014.

Alan M Leslie, Ori Friedman, and Tim P German. Core mechanisms in 'theory of mind'. *Trends in cognitive sciences*, 8(12):528–533, 2004.

Bob Lew, Jenny Huen, Pengpeng Yu, Lu Yuan, Dong-Fang Wang, Fan Ping, Mansor Abu Talib, David Lester, and Cun-Xian Jia. Associations between depression, anxiety, stress, hopelessness, subjective well-being, coping styles and suicide in Chinese university students. *PLOS ONE*, 14(7):e0217372, July 2019. ISSN 1932-6203. doi:10.1371/journal.pone.0217372. URL `https://dx.plos.org/10.1371/journal.pone.0217372`.

Glyn Lewis and Margaret Booth. Are cities bad for your mental health? *Psychological Medicine*, 24(4):913–916, 1994.

Shuhong Lin, Boaz Keysar, and Nicholas Epley. Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology*, 46(3):551–556, 2010.

Karina J Linnell, Serge Caparos, Jan W de Fockert, and Jules Davidoff. Urbanization decreases attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance*, 39(5):1232, 2013.

Karina J Linnell, Serge Caparos, and Jules Davidoff. Urbanization increases left-bias in line-bisection: an expression of elevated levels of intrinsic alertness? *Frontiers in psychology*, 5:1127, 2014.

Jose Lobo, Luis MA Bettencourt, Michael E Smith, and Scott Ortman. Settlement scaling theory: Bridging the study of ancient and contemporary urban systems. *Urban Studies*, 57(4):731–747, 2020.

Tom Loney and Nico J Nagelkerke. The individualistic fallacy, ecological studies and instrumental variables: a causal interpretation. *Emerging Themes in Epidemiology*, 11 (1):18, December 2014. ISSN 1742-7622. doi:10.1186/1742-7622-11-18. URL `https://ete-online.biomedcentral.com/articles/10.1186/1742-7622-11-18`.

Yu Luo and Jicong Zhang. The effect of tactile training on sustained attention in young adults. *Brain Sciences*, 10(10):695, 2020.

Antoine Lutz, Heleen A Slagter, Nancy B Rawlings, Andrew D Francis, Lawrence L Greischar, and Richard J Davidson. Mental training enhances attentional stability: neural and behavioral evidence. *Journal of Neuroscience*, 29(42):13418–13427, 2009.

Winnie WS Mak, Cecilia YM Poon, Loraine YK Pun, and Shu Fai Cheung. Meta-analysis of stigma and mental health. *Social science & medicine*, 65(2):245–261, 2007.

Scott Marek, Brenden Tervo-Clemmens, Finnegan J Calabro, David F Montez, Benjamin P Kay, Alexander S Hatoum, Meghan Rose Donohue, William Foran, Ryland L Miller, Timothy J Hendrickson, et al. Reproducible brain-wide association studies require thousands of individuals. *Nature*, 603(7902):654–660, 2022.

Alison Mary, Hichem Slama, Philippe Mousty, Isabelle Massat, Tatiana Capiau, Virginie Drabs, and Philippe Peigneux. Executive and attentional contributions to theory of mind deficit in attention deficit/hyperactivity disorder (adhd). *Child neuropsychology*, 22(3): 345–365, 2016.

John C. Mazziotta, Arthur W. Toga, Alan Evans, Peter Fox, and Jack Lancaster. A Probabilistic Atlas of the Human Brain: Theory and Rationale for Its Development: The International Consortium for Brain Mapping (ICBM). *NeuroImage*, 2(2, Part A):89–101, 1995. ISSN 1053-8119. doi:10.1006/nimg.1995.1012.

Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, pages 415–444, 2001.

Hygor Piaget M Melo, André A Moreira, Élcio Batista, Hernán A Makse, and José S Andrade. Statistical signs of social influence on suicides. *Scientific reports*, 4(1):1–6, 2014.

S. Milgram. The Experience of Living in Cities. *Science*, 167(3924):1461–1468, March 1970a. ISSN 0036-8075, 1095-9203. doi:10.1126/science.167.3924.1461. URL `http://www.sciencemag.org/cgi/doi/10.1126/science.167.3924.1461`.

Stanley Milgram. The experience of living in cities. *Science*, 167(3924):1461–1468, 1970b.

John Mirowsky and Catherine E Ross. Age and the effect of economic hardship on depression. *Journal of health and social behavior*, pages 132–150, 2001.

Lewis Mitchell, Morgan R Frank, Kameron Decker Harris, Peter Sheridan Dodds, and Christopher M Danforth. The geography of happiness: Connecting twitter sentiment and

expression, demographics, and objective characteristics of place. *PloS one*, 8(5):e64417, 2013.

Carlos Molinero and Stefan Thurner. How the geometry of cities determines urban scaling laws. *Journal of the Royal Society interface*, 18(176):20200705, 2021.

Gerald Mollenhorst, Beate Völker, and Henk Flap. Social contexts and core discussion networks: Using a choice approach to study similarity in intimate relationships. *Social Forces*, 86(3):937–965, 2008.

Kelly A Mollica, Barbara Gray, and Linda K Trevino. Racial homophily and its persistence in newcomers' social networks. *Organization Science*, 14(2):123–136, 2003.

NC Moore. The personality and mental health of flat dwellers. *The British Journal of Psychiatry*, 128(3):259–261, 1976.

Philip S. Morrison and Mikko Weckroth. Human values, subjective well-being and the metropolitan region. *Regional Studies*, 52(3):325–337, March 2018. ISSN 0034-3404, 1360-0591. doi:10.1080/00343404.2017.1331036. URL `https://www.tandfonline.com/doi/full/10.1080/00343404.2017.1331036`.

Kostas Mouratidis. Compact city, urban sprawl, and subjective well-being. *Cities*, 92:261–272, September 2019. ISSN 02642751. doi:10.1016/j.cities.2019.04.013. URL `https://linkinghub.elsevier.com/retrieve/pii/S0264275118316184`.

Jaap MJ Murre and Joeri Dros. Replication and analysis of ebbinghaus' forgetting curve. *PloS one*, 10(7):e0120644, 2015.

Elizbar A Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9 (1):141–142, 1964.

Anthony Nardone, Joey Chiang, and Jason Corburn. Historic redlining and urban health today in us cities. *Environmental Justice*, 13(4):109–119, 2020.

MEJ Newman. Resource letter cs–1: Complex systems. *American Journal of Physics*, 79 (8):800–810, 2011.

Brian Nolan, Max Roser, and Stefan Thewissen. Gdp per capita versus median household income: What gives rise to the divergence over time and how does this vary across oecd countries? *Review of Income and Wealth*, 65(3):465–494, 2019.

Sébastien Normand, Barry H Schneider, Matthew D Lee, Marie-France Maisonneuve, Angelina Chupetlovska-Anastasova, Sally M Kuehn, and Philippe Robaey. Continuities and changes in the friendships of children with and without adhd: A longitudinal, observational study. *Journal of Abnormal Child Psychology*, 41:1161–1175, 2013.

Office of Management and Budget. Recommendations from the metropolitan and micropolitan statistical area standards review committee to the office of management and budget concerning changes to the 2010 standards for delineating metropolitan and micropolitan statistical areas. *Fed. Regist.*, 86:5263–5266, 2021.

University of Wisconsin School of Medicine and Public Health. Area deprivation index 2019, 2019.

Janet Okamoto, C Anderson Johnson, Adam Leventhal, Joel Milam, Mary Ann Pentz, David Schwartz, and Thomas W Valente. Social network status and depression among adolescents: An examination of social network influences and depressive symptoms in a chinese sample. *Research in human development*, 8(1):67–88, 2011.

Marcos Oliveira, Carmelo Bastos-Filho, and Ronaldo Menezes. The scaling of crime concentration in cities. *PloS one*, 12(8):e0183110, 2017.

Marcos Oliveira, Fariba Karimi, Maria Zens, Johann Schaible, Mathieu Génois, and Markus Strohmaier. Group mixing drives inequality in face-to-face gatherings. *Communications Physics*, 5(1):1–9, 2022.

Scott G. Ortman, Andrew H. F. Cabaniss, Jennie O. Sturm, and Luís M. A. Bettencourt. The pre-history of urban scaling. *PLOS ONE*, 9(2):1–10, 02 2014a. doi:10.1371/journal.pone.0087902. URL https://doi.org/10.1371/journal.pone.0087902.

Scott G Ortman, Andrew HF Cabaniss, Jennie O Sturm, and Luís MA Bettencourt. The pre-history of urban scaling. *PloS one*, 9(2):e87902, 2014b.

Scott G. Ortman, Andrew H. F. Cabaniss, Jennie O. Sturm, and Luís M. A. Bettencourt. Settlement scaling and increasing returns in an ancient society. *Science Advances*, 1(1), 2015. doi:10.1126/sciadv.1400066. URL https://advances.sciencemag.org/content/1/1/e1400066.

Scott G. Ortman, Kaitlyn E. Davis, José Lobo, Michael E. Smith, Luis M.A. Bettencourt, and Aaron Trumbo. Settlement scaling and economic change in the central andes. *Journal of Archaeological Science*, 73:94 – 106, 2016. ISSN 0305-4403. doi:https://doi.org/10.1016/j.jas.2016.07.012. URL http://www.sciencedirect.com/science/article/pii/S0305440316300991.

Scott G Ortman, José Lobo, and Michael E Smith. Cities: Complexity, theory and history. *Plos one*, 15(12):e0243621, 2020.

Andrew J. Oswald and Stephen Wu. Well-Being across America. *Review of Economics and Statistics*, 93(4):1118–1134, November 2011. ISSN 0034-6535, 1530-9142. doi:10.1162/REST_a_00133. URL http://www.mitpressjournals.org/doi/10.1162/REST_a_00133.

Vanessa Panaite, Jonathan Rottenberg, and Lauren M. Bylsma. Daily Affective Dynamics Predict Depression Symptom Trajectories Among Adults with Major and Minor Depression. *Affective Science*, 1(3):186–198, September 2020. ISSN 2662-2041, 2662-205X. doi:10.1007/s42761-020-00014-w. URL `http://link.springer.com/10.1007/s42761-020-00014-w`.

Filippo Passetti, Yogita Chudasama, and Trevor W Robbins. The frontal cortex of the rat and visual attentional performance: dissociable functions of distinct medial prefrontal subregions. *Cerebral cortex*, 12(12):1254–1268, 2002.

B Keith Payne and Jason W Hannay. Implicit bias reflects systemic racism. *Trends in cognitive sciences*, 25(11):927–936, 2021.

B Keith Payne and Heidi A Vuletich. Policy insights from advances in implicit bias research. *Policy Insights from the Behavioral and Brain Sciences*, 5(1):49–56, 2018.

B Keith Payne, Heidi A Vuletich, and Kristjen B Lundberg. The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, 28(4):233–248, 2017.

B Keith Payne, Heidi A Vuletich, and Jazmin L Brown-Iannuzzi. Historical roots of implicit bias in slavery. *Proceedings of the National Academy of Sciences*, 116(24):11693–11698, 2019.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Michael Peer, Mordechai Hayman, Bar Tamir, and Shahar Arzy. Brain coding of social network structure. *Journal of Neuroscience*, 41(22):4897–4909, 2021.

Thomas F Pettigrew, Oliver Christ, Ulrich Wagner, and Jost Stellmacher. Direct and indirect intergroup contact effects on prejudice: A normative interpretation. *International Journal of intercultural relations*, 31(4):411–425, 2007.

Jonathan D Power, Alexander L Cohen, Steven M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, Bradley L Schlaggar, et al. Functional network organization of the human brain. *Neuron*, 72(4): 665–678, 2011.

Naomi Priest, Yin Paradies, Angeline Ferdinand, Lobna Rouhani, and Margaret Kelaher. Patterns of intergroup contact in public spaces: Micro-ecology of segregation in australian communities. *Societies*, 4(1):30–44, 2014.

Antoinette Prouteau, Hélène Verdoux, Catherine Briand, Alain Lesage, Pierre Lalonde, Luc Nicole, Daniel Reinharz, and Emmanuel Stip. The crucial role of sustained attention in community functioning in outpatients with schizophrenia. *Psychiatry Research*, 129(2): 171–177, 2004.

James R Rae, Anna-Kaisa Newheiser, and Kristina R Olson. Exposure to racial out-groups and implicit race bias in the united states. *Social Psychological and Personality Science*, 6(5):535–543, 2015.

Karthik Rajkumar, Guillaume Saint-Jacques, Iavor Bojinov, Erik Brynjolfsson, and Sinan Aral. A causal test of the strength of weak ties. *Science*, 377(6612):1304–1310, 2022.

Nornadiah Mohd Razali, Yap Bee Wah, et al. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*, 2(1):21–33, 2011.

Rachel A Razza, Anne Martin, and Jeanne Brooks-Gunn. Associations among family environment, sustained attention, and school readiness for low-income children. *Developmental psychology*, 46(6):1528, 2010.

L Song Richardson. Implicit racial bias and racial anxiety: Implications for stops and frisks. *Ohio St. J. Crim. L.*, 15:73, 2017.

Charles E Rosenberg. *The cholera years: The United States in 1832, 1849, and 1866*. University of Chicago Press, 2009.

Monica D Rosenberg, Emily S Finn, Dustin Scheinost, Xenophon Papademetris, Xilin Shen, R Todd Constable, and Marvin M Chun. A neuromarker of sustained attention from whole-brain functional connectivity. *Nature neuroscience*, 19(1):165–171, 2016a.

Monica D Rosenberg, Sheng Zhang, Wei-Ting Hsu, Dustin Scheinost, Emily S Finn, Xilin Shen, R Todd Constable, Chiang-Shan R Li, and Marvin M Chun. Methylphenidate modulates functional network connectivity to enhance attention. *Journal of Neuroscience*, 36(37):9547–9557, 2016b.

Monica D Rosenberg, Emily S Finn, Dustin Scheinost, Robert T Constable, and Marvin M Chun. Characterizing attention with predictive network models. *Trends in cognitive sciences*, 21(4):290–302, 2017.

Monica D Rosenberg, Wei-Ting Hsu, Dustin Scheinost, R Todd Constable, and Marvin M Chun. Connectome-based models predict separable components of attention in novel individuals. *Journal of cognitive neuroscience*, 30(2):160–173, 2018.

Monica D Rosenberg, Dustin Scheinost, Abigail S Greene, Emily W Avery, Young Hye Kwon, Emily S Finn, Ramachandran Ramani, Maolin Qiu, R Todd Constable, and Marvin M Chun. Functional connectivity predicts changes in attention observed across minutes, days, and months. *Proceedings of the National Academy of Sciences*, 117(7):3797–3807, 2020.

J Niels Rosenquist, James H Fowler, and Nicholas A Christakis. Social network determinants of depression. *Molecular psychiatry*, 16(3):273–281, 2011.

Andrew F Rossi, Luiz Pessoa, Robert Desimone, and Leslie G Ungerleider. The prefrontal cortex and the executive control of attention. *Experimental brain research*, 192:489–497, 2009.

Marcel Salathé, Maria Kazandjieva, Jung Woo Lee, Philip Levis, Marcus W Feldman, and James H Jones. A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences*, 107(51):22020–22025, 2010.

Horacio Samaniego, Mauricio Franco-Cisterna, and Boris Sotomayor-Gómez. The topology of communicating across cities of increasing sizes, or the complex task of "reaching out" in larger cities. *Theories and Models of Urbanization: Geography, Economics and Computing Sciences*, pages 97–118, 2020.

James Saxon. The local structures of human mobility in Chicago. *Environment and Planning B: Urban Analytics and City Science*, page 239980832094953, August 2020. ISSN 2399-8083, 2399-8091. doi:10.1177/2399808320949539. URL `http://journals.sagepub.com/doi/10.1177/2399808320949539`.

Thomas C Schelling. Dynamic models of segregation. *Journal of mathematical sociology*, 1 (2):143–186, 1971.

Thomas C Schelling. *Micromotives and macrobehavior*. WW Norton & Company, 2006.

Kathryn E Schertz, Sonya Sachdeva, Omid Kardan, Hiroki P Kotabe, Kathleen L Wolf, and Marc G Berman. A thought in the park: The influence of naturalness and low-level visual features on expressed thoughts. *Cognition*, 174:82–93, 2018.

Markus Schläpfer, Luís MA Bettencourt, Sébastian Grauwin, Mathias Raschke, Rob Claxton, Zbigniew Smoreda, Geoffrey B West, and Carlo Ratti. The scaling of human interactions with city size. *Journal of the Royal Society Interface*, 11(98):20130789, 2014a.

Markus Schläpfer, Luís MA Bettencourt, Sébastian Grauwin, Mathias Raschke, Rob Claxton, Zbigniew Smoreda, Geoffrey B West, and Carlo Ratti. The scaling of human interactions with city size. *Journal of the Royal Society Interface*, 11(98):20130789, 2014b.

Ralf Schmälzle, Matthew Brook O'Donnell, Javier O Garcia, Christopher N Cascio, Joseph Bayer, Danielle S Bassett, Jean M Vettel, and Emily B Falk. Brain connectivity dynamics during social interaction reflect social network structure. *Proceedings of the National Academy of Sciences*, 114(20):5153–5158, 2017.

Konrad Schnabel, Jens B Asendorpf, and Anthony G Greenwald. Assessment of individual differences in implicit cognition: A review of iat measures. *European Journal of Psychological Assessment*, 24(4):210, 2008.

Brian J Scholl. What have we learned about attention from multiple object tracking (and vice versa). *Computation, cognition, and Pylyshyn*, pages 49–78, 2009.

Christopher R Sears and Zenon W Pylyshyn. Multiple object tracking and attentional processing. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 54(1):1, 2000.

SuGyun Seo, JeeHye Jeon, YoungSook Chong, and JeongShin An. The Relations Among Relatedness Needs, Subjective Well-Being, and Depression of Korean Elderly. *Journal of Women & Aging*, 27(1):17–34, January 2015. ISSN 0895-2841, 1540-7322. doi:10.1080/08952841.2014.929406. URL https://www.tandfonline.com/doi/full/10.1080/08952841.2014.929406.

Xilin Shen, Xenophon Papademetris, and R Todd Constable. Graph-theory based parcellation of functional subunits in the brain from resting-state fmri data. *Neuroimage*, 50(3): 1027–1035, 2010.

Feng Sheng, Yi Liu, Bin Zhou, Wen Zhou, and Shihui Han. Oxytocin modulates the racial bias in neural responses to others' suffering. *Biological Psychology*, 92(2):380–386, 2013.

Georg Simmel. *The sociology of georg simmel*, volume 92892. Simon and Schuster, 1950.

Georg Simmel. The metropolis and mental life. In *The urban sociology reader*, pages 37–45. Routledge, 2012.

Archana Singh and Nishi Misra. Loneliness, depression and sociability in old age. *Industrial psychiatry journal*, 18(1):51, 2009.

John Skvoretz. Diversity, integration, and social ties: Attraction versus repulsion as drivers of intra-and intergroup relations. *American Journal of Sociology*, 119(2):486–517, 2013.

S. A. Stansfeld, J. Head, and M. G. Marmot. Explaining social class differences in depression and well-being. *Social Psychiatry and Psychiatric Epidemiology*, 33(1):1–9, December 1997. ISSN 0933-7954, 1433-9285. doi:10.1007/s001270050014. URL http://link.springer.com/10.1007/s001270050014.

Andrew Steptoe, Samantha Dockray, and Jane Wardle. Positive affect and psychobiological processes relevant to health. *Journal of Personality*, 77(6):1747–1776, 2009. doi:https://doi.org/10.1111/j.1467-6494.2009.00599.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-6494.2009.00599.x.

Andrew Steptoe, Angus Deaton, and Arthur A Stone. Subjective wellbeing, health, and ageing. *The Lancet*, 385(9968):640–648, February 2015. ISSN 01406736. doi:10.1016/S0140-6736(13)61489-0. URL https://linkinghub.elsevier.com/retrieve/pii/S0140673613614890.

Andrew Stier, Marc Berman, and Luis Bettencourt. Covid-19 attack rate increases with city size. *Mansueto Institute for Urban Innovation Research Paper Forthcoming*, 2020.

Andrew Stier, Sina Sajjadi, Fariba Karimi, Luis Bettencourt, and Marc G Berman. City population, majority group size, and residential segregation drive implicit racial biases in us cities. *Majority Group Size, and Residential Segregation Drive Implicit Racial Biases in US Cities (January 27, 2023)*, 2023.

Andrew J Stier, Kathryn E Schertz, Nak Won Rim, Carlos Cardenas-Iniguez, Benjamin B Lahey, Luís MA Bettencourt, and Marc G Berman. Evidence and theory for lower rates of depression in larger us urban areas. *Proceedings of the National Academy of Sciences*, 118(31):e2022472118, 2021.

Andrew J Stier, Sina Sajjadi, Luis Bettencourt, Fariba Karimi, and Marc G Berman. Effects of racial segregation on economic productivity in us cities. *arXiv preprint arXiv:2212.03147*, 2022a.

Andrew J Stier, Sina Sajjadi, Luís M A Bettencourt, Fariba Karimi, and Marc G Berman. Effects of racial segregation on economic productivity in u.s. cities. *In Preparation*, 2022b.

Andrew J. Stier, Sina Sajjadi, Luis M. A. Bettencourt, Fariba Karimi, and Marc G. Berman. Effects of racial segregation on economic productivity in u.s. cities. *arXiv*, 2022c. doi:10.48550/ARXIV.2212.03147. URL `https://arxiv.org/abs/2212.03147`.

Andrew J Stier, Kathryn E Schertz, Nak Won Rim, Carlos Cardenas-Iniguez, Benjamin B Lahey, Luís MA Bettencourt, and Marc G Berman. Reply to huth et al.: Cities are defined by their spatially aggregated socioeconomic networks. *Proceedings of the National Academy of Sciences*, 119(2):e2119313118, 2022d.

James Stiller and Robin IM Dunbar. Perspective-taking and memory capacity predict social network size. *Social Networks*, 29(1):93–104, 2007.

Rebecca Storey. An estimate of mortality in a pre-columbian urban population. *American Anthropologist*, 87(3):519–535, 1985.

Katherine R Storrs, Seyed-Mahdi Khaligh-Razavi, and Nikolaus Kriegeskorte. Noise ceiling on the crossvalidated performance of reweighted models of representational dissimilarity: Addendum to khaligh-razavi & kriegeskorte (2014). *BioRxiv*, 2020.

Kristina Sundquist, Gölin Frank, and Jan Sundquist. Urbanisation and incidence of psychosis and depression: Follow-up study of 4.4 million women and men in Sweden. *British Journal of Psychiatry*, 184(4):293–298, April 2004. ISSN 0007-1250, 1472-1465. doi:10.1192/bjp.184.4.293. URL `https://www.cambridge.org/core/product/identifier/S0007125000077722/type/journal_article`.

Hanna Swaab-Barneveld, Leo De Sonneville, Peggy Cohen-Kettenis, Anneke Gielen, Jan Buitelaar, and Herman Van Engeland. Visual sustained attention in a child psychiatric population. *Journal of the American Academy of Child & Adolescent Psychiatry*, 39(5): 651–659, 2000.

Tiit Tammaru, Magnus Strömgren, Maarten Van Ham, and Alexander M Danzer. Relations between residential and workplace segregation among newly arrived immigrant men and women. *Cities*, 59:131–138, 2016.

Zeynep Baran Tatar and Alparslan Cansız. Executive function deficits contribute to poor theory of mind abilities in adults with adhd. *Applied Neuropsychology: Adult*, 29(2): 244–251, 2022.

Sylvia Terbeck, Guy Kahane, Sarah McTavish, Julian Savulescu, Philip J Cowen, and Miles Hewstone. Propranolol reduces implicit negative racial bias. *Psychopharmacology*, 222(3): 419–424, 2012.

Mike Thelwall. Homophily in myspace. *Journal of the American Society for Information Science and Technology*, 60(2):219–231, 2009.

Andromachi Tseloni, Jen Mailley, Graham Farrell, and Nick Tilley. Exploring the international decline in crime rates. *European Journal of Criminology*, 7(5):375–394, 2010.

Riley Tucker, Daniel T O'Brien, Alexandra Ciomek, Edgar Castro, Qi Wang, and Nolan Edward Phillips. Who 'tweets' where and when, and how does it help understand crime rates at places? measuring the presence of tourists and commuters in ambient populations. *Journal of Quantitative Criminology*, 37(2):333–359, 2021.

N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee. N4itk: Improved n3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6): 1310–1320, 2010. ISSN 0278-0062. doi:10.1109/TMI.2010.2046908.

Roger S Ulrich. Biophilia, biophobia, and natural landscapes. *The biophilia hypothesis*, 7: 73–137, 1993.

Department of Economic United Nations and Population Division Social Affairs. World urbanization prospects: The 2018 revision (st/esa/ser.a/420). new york: United nations. *United Nations World Urbanization Prospects*, 2019.

Bob Van der Zwaan and Ari Rabl. Prospects for pv: a learning curve analysis. *Solar energy*, 74(1):19–31, 2003.

Dianne A Van Hemert, Fons JR Van De Vijver, and Ype H Poortinga. The beck depression inventory as a measure of subjective well-being: A cross-national study. *Journal of Happiness Studies*, 3:257–286, 2002a.

Dianne A Van Hemert, Fons JR Van De Vijver, and Ype H Poortinga. The beck depression inventory as a measure of subjective well-being: A cross-national study. *Journal of Happiness Studies*, 3(3):257–286, 2002b.

Christiaan H Vinkers, Marian Joëls, Yuri Milaneschi, René S Kahn, Brenda WJH Penninx, and Marco PM Boks. Stress exposure across the life span cumulatively increases depression risk and is moderated by neuroticism. *Depression and Anxiety*, 31(9):737–745, 2014.

Heidi A Vuletich and B Keith Payne. Stability and change in implicit bias. *Psychological science*, 30(6):854–862, 2019.

Ulrich Wagner, Rolf Van Dick, Thomas F Pettigrew, and Oliver Christ. Ethnic prejudice in east and west germany: The explanatory power of intergroup contact. *Group Processes & Intergroup Relations*, 6(1):22–36, 2003.

Ulrich Wagner, Oliver Christ, Thomas F Pettigrew, Jost Stellmacher, and Carina Wolf. Prejudice and minority proportion: Contact instead of threat effects. *Social psychology quarterly*, 69(4):380–390, 2006.

Geoffrey S Watson. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 359–372, 1964.

Lilian Weng, Márton Karsai, Nicola Perra, Filippo Menczer, and Alessandro Flammini. Attention on weak ties in social and communication networks. *Complex spreading phenomena in social Systems: Influence and contagion in real-world social networks*, pages 213–228, 2018.

Aliza Werner-Seidler, Mohammad H Afzali, Cath Chapman, Matthew Sunderland, and Tim Slade. The relationship between social support networks and depression in the 2007 national survey of mental health and well-being. *Social psychiatry and psychiatric epidemiology*, 52(12):1463–1473, 2017.

Holly White and Priti Shah. Focus: Attention science: Attention in urban and natural environments. *The Yale journal of biology and medicine*, 92(1):115, 2019.

Michael J White. Segregation and diversity measures in population distribution. *Population index*, pages 198–221, 1986.

Edward O Wilson. *Biophilia*. Harvard university press, 1986.

Thomas C Wilson. Urbanism and tolerance: A test of some hypotheses drawn from wirth and stouffer. *American Sociological Review*, pages 117–123, 1985.

Louis Wirth. Urbanism as a way of life. *American journal of sociology*, 44(1):1–24, 1938.

Piotr A Woźniak, Edward J Gorzelańczyk, and Janusz A Murakowski. Two components of long-term memory. *Acta neurobiologiae experimentalis*, 55(4):301–305, 1995.

Esther XW Wu, Gwenisha J Liaw, Rui Zhe Goh, Tiffany TY Chia, Alisia MJ Chee, Takashi Obana, Monica D Rosenberg, BT Thomas Yeo, and Christopher L Asplund. Overlapping attentional networks yield divergent behavioral predictions across tasks: neuromarkers for diffuse and focused attention? *NeuroImage*, 209:116535, 2020.

Kaiyuan Xu, Brian Nosek, and Anthony Greenwald. Psychology data from the race implicit association test on the project implicit demo website. *Journal of Open Psychology Data*, 2(1), 2014.

Yang Xu, Alexander Belyi, Paolo Santi, and Carlo Ratti. Quantifying segregation in an integrated urban physical-social space. *Journal of the Royal Society Interface*, 16(160): 20190536, 2019.

Amir Hossein Yazdavar, Hussein S Al-Olimat, Monireh Ebrahimi, Goonmeet Bajaj, Tanvi Banerjee, Krishnaprasad Thirunarayan, Jyotishman Pathak, and Amit Sheth. Semi-supervised approach to monitoring clinical depressive symptoms in social media. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, pages 1191–1198, 2017.

Kwangsun Yoo, Monica D Rosenberg, Young Hye Kwon, Qi Lin, Emily W Avery, Dustin Sheinost, R Todd Constable, and Marvin M Chun. A brain-based general measure of attention. *Nature human behaviour*, 6(6):782–795, 2022.

Jonathan C Ziegert and Paul J Hanges. Employment discrimination: the role of implicit attitudes, motivation, and a climate for racial bias. *Journal of applied psychology*, 90(3): 553, 2005.