Computers, Environment and Urban Systems 113 (2024) 102173

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

The great equalizer? Mixed effects of social infrastructure on diverse encounters in cities

Timothy Fraser [a,*,1], Takahiro Yabe [b,c,*,1], Daniel P. Aldrich [d], Esteban Moro [e,f]

[a] Systems Engineering Program, Cornell University, Ithaca, NY 14850, USA
[b] Institute of Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[c] Department of Technology Management and Innovation and Center for Urban Science and Progress, Tandon School of Engineering, New York University, Brooklyn, NY 11201, USA
[d] School of Public Policy and Urban Affairs, Northeastern University, Boston, MA 02115, USA
[e] Network Science Institute, Department of Physics, Northeastern University, Boston, MA, 02115, USA
[f] MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

ARTICLE INFO

Keywords:
Social infrastructure
Diversity
Human mobility
Social capital

ABSTRACT

Casual encounters with diverse groups of people in urban spaces have been shown to foster social capital and trust, leading to higher quality of life, civic participation, and community resilience to hazards. To promote such diverse encounters and cultivate social ties, policymakers develop social infrastructure sites, such as community centers, parks, and plazas. However, their effects on the diversity of encounters, compared to baseline sites (e.g., grocery stores), have not been fully understood. In this study, we use a large-scale, privacy-enhanced mobility dataset of >120 K anonymized mobile phone users in the Boston area to evaluate the effects of social infrastructure sites on the observed frequencies of inter-income and inter-race encounters. Contrary to our intuition that all social infrastructure sites promote diverse encounters, we find the effects to be mixed and more nuanced. Overall, parks and social businesses promote more inter-income encounters, while community spaces promote more same-income encounters, but each produces opposite effects for inter-race encounters. Parks and community spaces located in low-income neighborhoods were shown to result in higher inter-income and inter-race encounters compared to ordinary sites, respectively, however, their associations were insignificant in high-income areas. These empirical results suggest that the type of social infrastructure and neighborhood traits may alter levels of diverse encounters.

1. Introduction

Since Robert Putnam's landmark study "Bowling Alone" (Putnam, 2000), a wealth of studies have documented how social capital - the social ties that build trust and reciprocity among residents - boosts the health (Kawachi et al., 2008), quality of life (Rogers et al., 2011), civic participation (Johnson, 2010), and resilience to hazards (Aldrich, 2012) in urban neighborhoods (see (Aldrich and Meyer, 2015; Villalonga-Olives et al., 2018) for comprehensive reviews). The importance of neighborhood elements that promote diverse encounters and interactions, including density and mixed use spaces, was suggested by Jane Jacobs in her seminal book "The Death and Life of Great American Cities" (Jacobs, 1993). Fortunately, individuals' social capital is not static but can change. Social infrastructure, referring to sites like community centers, plazas, and bookstores, is thought to help build social ties, by creating public spaces for friendly social encounters (Klinenberg, 2018). Social infrastructure comes in multiple types (Latham and Layton, 2019), including parks (in this study, green space, community gardens, fountains, and plazas), community spaces (libraries, community centers, and public meeting spaces), and social businesses (cafes and bookstores), among others (e.g., places of worship, which this study cannot examine due to data privacy restrictions) (Fraser et al., 2022). More socially-friendly urban design has been linked with more social connectivity (Hanibuchi et al., 2012), place attachment (Kim and Kaplan, 2004), frequent unplanned social interaction (Lund, 2003; Wood et al., 2010), political participation (Leyden, 2003), and group

* Corresponding authors.
E-mail addresses: timothy.fraser.1@gmail.com (T. Fraser), takahiroyabe@nyu.edu (T. Yabe), daniel.aldrich@gmail.com (D.P. Aldrich), emoro@mit.edu (E. Moro).
[1] Fraser and Yabe are equal contribution authors.

https://doi.org/10.1016/j.compenvurbsys.2024.102173
Received 28 February 2023; Received in revised form 2 July 2024; Accepted 11 August 2024
Available online 20 August 2024

participation (Podobnik, 2011).

Casual encounters at the dog park, the library, or cafe might not build close ties, but we would expect these places to foster short encounters and weak ties, but these are not inconsequential (Small, 2017); weak ties can build familiarity, neighborliness, trust, and reciprocity across group lines (Aldrich, 2012), each vital to democracy (Coleman, 1988; Putnam, 2000; Szreter and Woolcock, 2004). In an equitable society, we would hope that social infrastructure creates social encounters between residents from different backgrounds - a phenomenon known as *social mixing* - but this is not certain. Urban gentrification, income inequality, and racial stratification common in American cities likely impede many of these diverse encounters today, leading to inequitable social outcomes (Chetty et al., 2019; Hwang and Ding, 2020).

Fortunately, scholars have new tools available for measuring social encounters. Large-scale GPS location data collected from smartphone devices ('mobility data') enables us to observe the visitation patterns of individuals at an extremely high spatial and temporal granularity across a long period of time (Blondel et al., 2015). Mobility data has been used to answer various questions about individual human behavior (Gonzalez et al., 2008), for measuring migration patterns (Blumenstock, 2012), for predicting epidemic spreading dynamics (Aleta et al., 2020), and for disaster risk management and resilience (Yabe et al., 2022). A recent study has used mobility data to quantify the segregation that occurs among urban physical encounters at various places (points-of-interest; POIs), revealing that 55% of the income segregation we experience is due to our mobility behavior, not where we reside (Moro et al., 2021). A more in-depth investigation into how implementing various types of places could improve the diversity of encounters in cities could connect such insights to concrete policy actions.

Consequently, we pose a question: How do our social encounters differ at social infrastructure sites, such as parks, libraries, or cafes, than we would experience at other ordinary urban spaces? In particular, we investigate two types of mixing in social encounters particularly salient in modern American cities: (1) inter-income encounters, where at specific sites, people might tend to encounter others from different income strata more than the same income strata. Alternatively, (2) inter-race encounters describe how certain sites might tend to bring people together from different racial backgrounds, more so than from the same racial background. In this study, we test three major hypotheses that collectively provide quantitative insights into the effects of social infrastructure for achieving a more diverse urban environment:

- **H1. Effects of social infrastructure on the diversity of encounters:** Social infrastructure is associated with higher **(H1.1)** inter-income and **(H1.2)** inter-race encounters, compared to ordinary sites. This hypothesis is supported by smaller-scale and qualitative studies (e.g., (Fraser and Naquin, 2022; Latham and Layton, 2019)), but is yet to be specifically tested quantitatively and at the urban scale.
- **H2. Heterogeneity across different social infrastructure types:** We expect community spaces and parks to promote more inter-group encounters, while social businesses may promote more same-group encounters, due to gentrification. We expect social businesses, community spaces, and parks to promote inter-race/inter-income encounters at different levels compared to each other.
- **H3. Interaction effects between infrastructure and characteristics of local area:** Social infrastructure produces mixed effects on diverse encounters, increasing inter-income encounters in more racially homogeneous places with fewer inter-race encounters, and vice versa.

Panels A and **B** show the area of Boston and the social infrastructure sites located within the area. **Panel C** shows the percentage of the three types of social infrastructure in our dataset, along with baseline sites used in the analysis. **Panels D** and **E** show illustrations of how inter-income and inter-race encounters are measured using mobile phone data. **Panels F** and **G** present box plots of the estimated inter-income and inter-race encounter indices for the types of social infrastructure sites. While the differences between baseline sites are small, we aim to further understand the heterogenous effects of social infrastructure sites on diverse encounters.

2. Methods

To investigate relationships between social infrastructure and diverse encounters, we used a 4-stage process, including (1) collecting points of interest, (2) mobility data-based estimates of diverse encounters, (3) statistical models to analyze the levels of diverse encounters, and (4) statistical simulations using these empirical models.

2.1. Data collection and processing

2.1.1. Points of interest

First, this study analyzed a sample of 356 geolocated points of interest (POIs). To account for urban/rural differences and city-by-city differences, our sampled POIs all fall within 13 contiguous core Boston neighborhoods, focusing on urban neighborhoods rather than more suburban or mixed locales in adjoining municipalities like Cambridge, Somerville, Brighton, etc.

To capture all known social infrastructure sites in the study area, we used Fraser and colleagues' publicly available dataset of Boston social infrastructure; that study layered Boston with a fishnet grid, collecting up to 20 sites matching a keyword search, repeated across a grid of 2 square-kilometer cells for 10 keywords returning valid results (out of 19 keywords tested) (Fraser et al., 2022). Validation efforts (Fraser et al., 2022) found these data demonstrate expected correlations with local-level measures of social capital, including validated indices, civil society membership rates, and civic participation measures. Further, they demonstrated strong internal and measurement validity when compared against human-generated maps of social infrastructure (gathered from online queries) and ground-truthed maps of social infrastructure (gathered by on-the-ground site visits). We examined social infrastructure sites from 9 (out of 10) keyword searches, including: 98 community spaces (keywords: libraries, community centers, and city hall facilities), 184 parks (keywords: parks, fountains, squares, and gardens), and 74 social businesses (bookstores and cafes). We exclude places of worship (mosques, churches, synagogues, etc.), for which we are unable to estimate mobility due to privacy restrictions. We paired these data with a group of baseline sites, sourced from Moro and colleagues' recent study of Boston mobility, including 35 post offices and 63 grocery stores and supermarkets. All site data was publicly available on Google Maps.

2.1.2. Measuring diversity of encounters with mobility data

Next, for each POI, we relied on privacy-enhanced mobility data from the Spectus Social Impact Program (as described in the Results), quantifying the levels of visitation to each POI from a random sample of over 120 K cell phone users who were opted-in to anonymized data collection for research purposes during the study time frame (just prior to February 2020). We classified each user's user's 'home' census block as the census block group where they spent the majority of their time between 10 PM and 6 AM during the study period. Then, we calculated quantities of interest for each POI, including (1) total site visits (normalized 0 to 1 for user privacy), (2) median distance traveled to each POI (from users' home), and (3) levels of inter-income encounters and (4) inter-race encounters, using the measurement strategy presented in the Results. This project received IRB exemption to use this mobility data from the MIT IRB office.

To test the hypotheses, we draw on a case study of micro-mobility data in contiguous Boston neighborhoods. We examine mobility from a set of 356 social infrastructure sites cataloged in a recent study (Fraser et al., 2022), including 98 community spaces, 184 parks, and 74 social businesses, compared against a baseline of ordinary places with similarly high foot traffic, including 35 post offices and 63 grocery stores and supermarkets, as shown in Fig. 1A – C. Spectus Inc. provided anonymized, high-resolution location pings for over 120 K opted-in devices within our 13 contiguous Boston neighborhoods, over two months. Each individual's home neighborhood was inferred by the data provider based on their most common Census Block Group (CBG) between 10 pm and 6 am, and the median income of the corresponding CBG obtained from the American Community Survey (ACS) was assigned to the individual as a proxy of their economic status. The individuals are categorized into four equally sized income quantiles based on their estimated income. The robustness of the choice of number of quantile bins on the results of experienced segregation was tested in a prior study (Moro et al., 2021; Yabe et al., 2023). Similarly, the majority race of the home CBG obtained from the ACS ('White', 'Black', 'Asian', or 'mixed' if the majority proportion is <50%) was assigned to the individual as a proxy for their race. Stays longer than 5 min were attributed to places by searching for the nearest place within 100 m. To control for pandemic-related mobility changes, we examined just mobility prior to February 2020. To ensure our associations are not just due to noise, meaning sites with just a few visitors, we cut the sample at the 20th percentile (0.03) and modeled the 80% most frequented sites ($n = 364$) in the core Boston area.

Using the observed visitation patterns from mobility data and estimated income and race features, we measure the diversity of encounters at these locations. Applying a previously proposed metric of segregation (Moro et al., 2021), we measured how much encounters occur between members of the same income group, and separately, racial group, on scales from 0 to 1. We then flipped the scale, such that 1 indicates that all encounters were with members from different strata (more inter-race/inter-income encounters), while 0 indicates that all encounters were with members from the same strata (fewer inter-race/inter-income encounters). Fig. 1D and E show schematics of how these metrics are computed for a single place.

Aggregate statistics of inter-income and inter-race encounters, shown in Figs. 1F and G show slight differences in the distributions across types of places, where parks and social businesses have higher inter-income encounters on average compared to baseline sites, while social businesses have higher inter-race encounters compared to baseline sites. In the following analyses, we aim to obtain a more in-depth understanding of the observed variability across places.

2.2. Modeling

To test hypotheses, we constructed several statistical simulations, based on the best-fitting models from several statistical models for each of our outcomes. For inter-income encounters, see **Table S1**; for inter-race encounters, see **Table S2**. For each, we sequentially built the model (Model 10, **Tables S1 & S2**) that maximized (1) predictive power and (2) best controlled for theoretical drivers of inter-group encounters. We employ beta regression models, the appropriate method for beta-

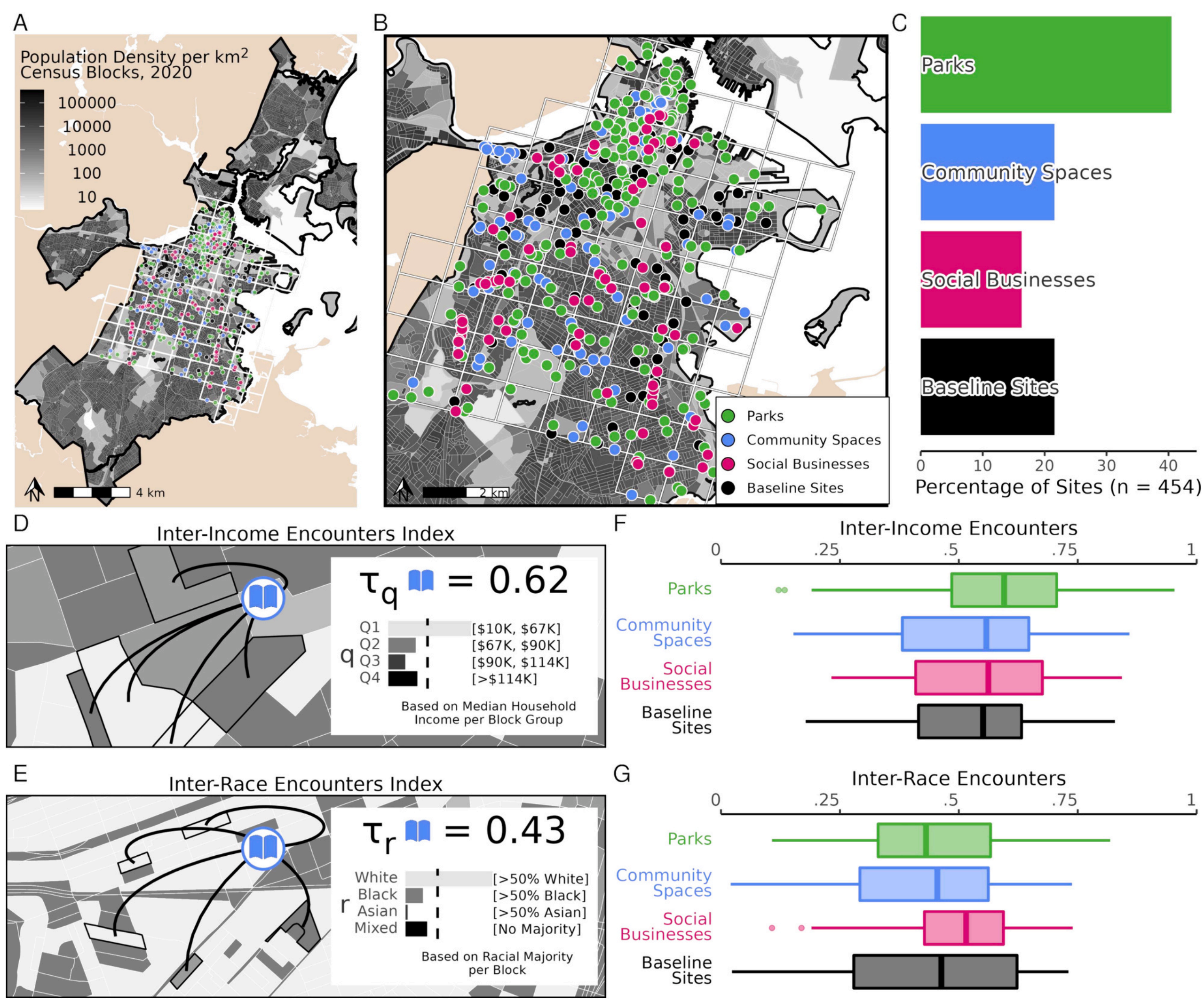

Fig. 1. Measuring inter-income and inter-race encounters at social infrastructure sites using mobility data.

distributed outcome variables ranging from 0 to 1, using the *betareg* package (v. 3.1–4) in R (v. 4.1.3) (Cribari-Neto and Zeileis, 2010). Data for area traits below sourced from the 2020 Decennial Census for census blocks where able; otherwise, we rely on American Community Survey 5-year averages (2016–2020) at the block group level using the *tidycensus* (v. 1.1) (Walker and Herman, 2021) and *censusapi* (v. 0.7.1) (Recht, 2020) packages in R.

2.2.1. Design of Interaction Effects

Below, we summarize our best-fitting models (see **SI Note 1.1** for a full discussion of models and controls. We recognize that inter-race and inter-income encounters are closely related, such that social infrastructure might have different effects on inter-income encounters not just independent of inter-race encounters, but depending on the level of inter-race encounters (and vice versa). For inter-income encounters, our best-fitting model cuts the sample into those POIs with "Lower", "Middle", or "Upper" levels of a mediating variable, namely inter-race encounters, using 0.33 and 0.66 as cutoffs. (4 sets of different cutoffs produced comparable results, and this model fit best; see **SI Note 1.1, 1.3**.) Then we regressed the interaction between this variable and each type of social infrastructure (parks, community spaces, and social businesses, vs. baseline sites); meanwhile, we controlled for the independent effects of inter-race encounters and each type of social infrastructure. When predicting inter-race encounters, we conversely applied interactions and controls for inter-income encounters as the mediating variable.

2.2.2. Control variables (in brief)

We controlled for 11 additional concepts, listed below and described extensively in the **Supplemental Information (SI)**. We controlled for mobility and user traits, including (1) the total normalized visit count per POI, (2) the median distance traveled to each POI, and (3) the population density of each POI's census block. We control for (4) user race using the percentage of users from a majority white block, (5) the percentage of users from block groups in the upper third to fourth income quartile; these ensure we are predicting homophily and heterophily in encounters, and not just race or income in disguise. Since some sites may simply have more inter-income/race encounters due to their demographic makeup and ease of access, we control for area traits, including (6) % white residents per census block, (7) median household income per block group, (8) % residents over age 65, distance in meters from each site to its nearest (9) metro stop and (10) bus stop, as well as (11) geographic differences using the census's 5 Public Use Microdata Areas (PUMAs). Finally, since most predictors held strong nonlinear associations with inter-income/race encounters, we applied second-degree polynomial transformations to all numeric variables, which afforded the greatest marginal improvement in model fit (R^2) while preserving low degree. Transformations improve model fit, allowing us to predict nearly 81% of inter-income encounters and 95% of inter-race encounters. (See **SI Note 1.1** for details on our modeling procedures, and **Tables S7 and S8** for descriptive statistics.)

2.2.3. Statistical simulations

Scholars have argued for over 20 years that predictions, not beta coefficients, are better tools for hypothesis testing, and are robust to model inference problems like multicollinearity and heteroskedasticity because predictions (particularly simulations) do not rely upon beta coefficients' standard errors (Greenberg and Parks, 1997; Imai et al., 2008; King et al., 2000). All simulations below were conducted in the **simulate** package (v. 0.0.0.9) in **R** (v. 4.1.3) (Fraser, 2022).

To test our hypotheses, we simulated marginal effects in Fig. 2, creating 32 marginal effects at the means from 1000 simulations each (16 per outcome), using the best fitting Model 10 in Tables S1 and S2. Each simulation held all variables at their median/modal observed values except for social infrastructure and the mediating variable, which were varied systematically to show all 24 combinations (2 outcomes × 4 POI types x 3 levels of mediating variable). By averaging these simulations over types of social infrastructure and levels of the mediating variable, we were able to compute the marginal percent change in inter-income/race encounters for (1) parks, (2) community spaces, (3) social businesses, and (4) social infrastructure overall, at each level of the mediating variable (Lower, Middle, Upper, and Overall), producing 32 effects (2 outcomes × 4 final types x 4 final levels). All estimates reflect the median marginal effect and upper and lower 95% confidence intervals from each set of 1000 simulations, with *p*-values showing the false positive rate (defined as the percentage of simulations on the opposite side of zero from the median marginal effect). We report exact simulations from Fig. 2 in **Tables S3 & S4**. (As robustness tests, we compared simulations from models with differing cutoffs for the mediating variable, providing largely consistent results; see **Tables S5 & S6**.)

For context, in Figs. 3 and 4, we conducted additional simulations for our 280 social infrastructure sites, using our best models to estimate the marginal percent change in inter-income/race encounters at each site due to being a park/community space/social business as opposed to a baseline site, holding all other POI traits were held at their real observed values. For each POI, we report the median marginal percent change

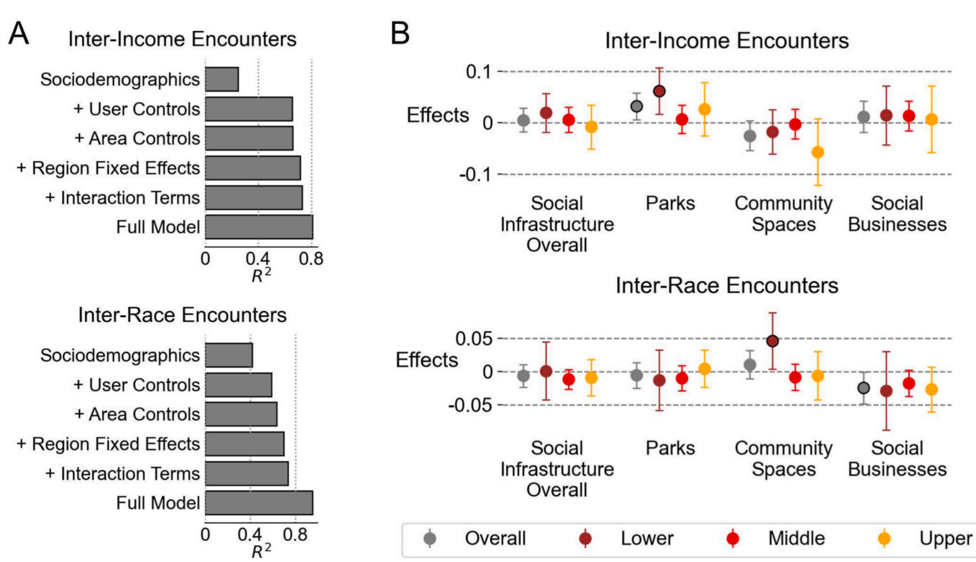

Fig. 2. Modeling the effects of social infrastructure on inter-income and inter-race encounters.

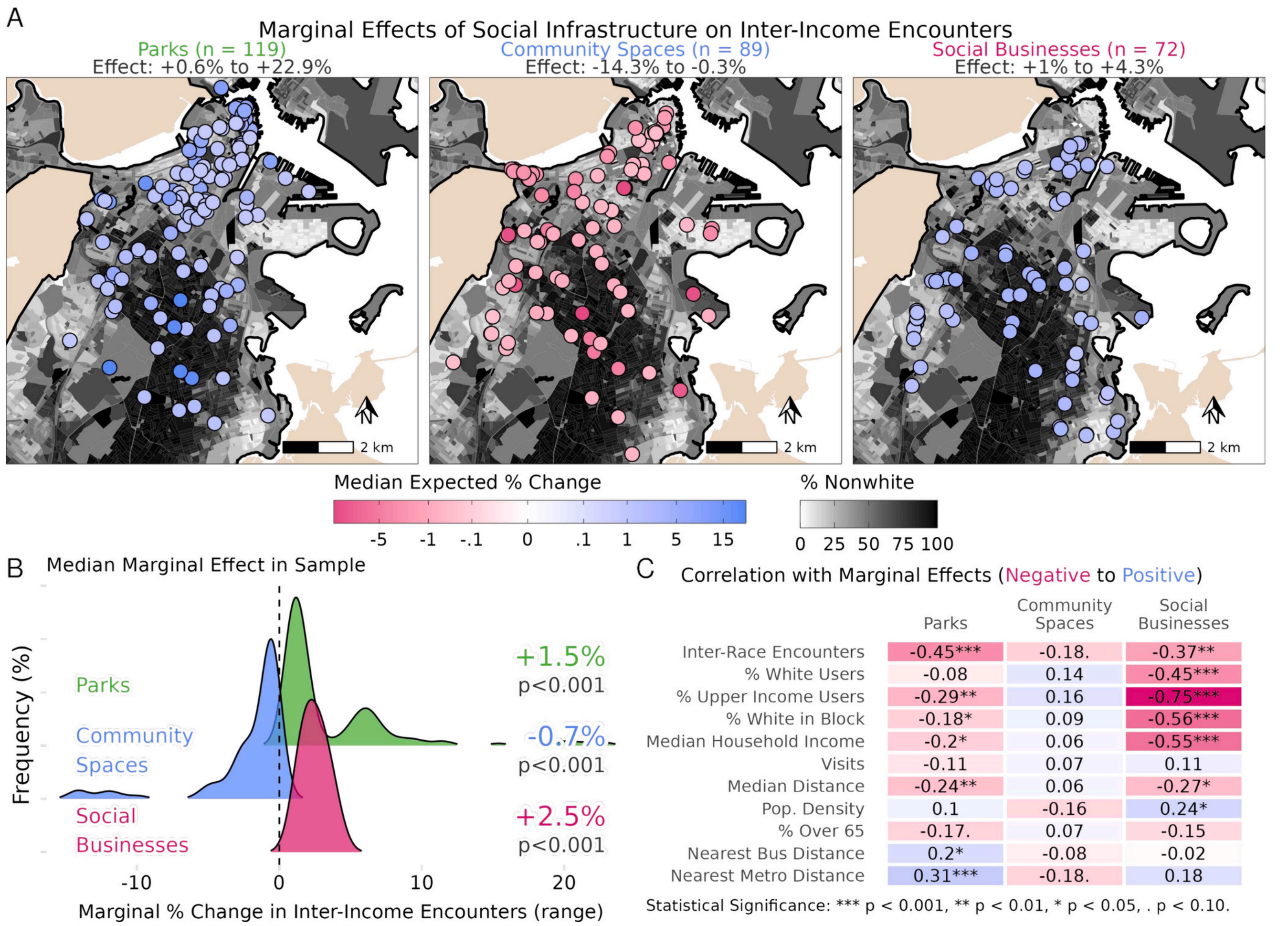

Fig. 3. Marginal effects of social infrastructure sites on inter-income encounters (compared to baseline).

from 1000 simulations, mapped in Figs. 3A and 4A, and then mapped the distributions of all 280 POI's median marginal percent changes attributed to each type of social infrastructure. Finally, we compared these median marginal effects against these POIs observed coefficients, to verify what kinds of parks, community spaces, and social businesses tended to see more/fewer diverse encounters than others.

3. Results

3.1. Mixed effects of social infrastructure on inter-race and inter-income encounters

To understand the variability in inter-income and inter-race encounters across places in our dataset, we built multiple regression models that use sociodemographic and economic characteristics of the places, and users, and located CBGs as exogenous variables. Our models narrow into 364 frequently visited sites, defined as within the upper 4 quintiles in terms of site visits, to ensure results are not skewed by rarely visited sites. Different groups of variables were added incrementally to assess the effects of the variables. Fig. 2A shows the R squared of the four levels of models to explain the variance in inter-income and inter-race encounters. For details about the model specifications and results, see **SI Note 1.1**.

Panel A shows the R^2 of the four levels of models to explain the variance in inter-income and inter-race encounters. Overall, the full model was able to explain 80% to 90% of the variance in diversity of encounters. **Panel B** shows the marginal effects of having different types of social infrastructure sites instead of baseline sites on inter-income and inter-race encounters. Results show *mixed effects*, where, for example, community spaces have negative effects on inter-income encounters but positive effects on inter-race encounters.

Sociodemographic variables which include the population density of the CBG that the places are located in, accounted for around 25% and 40% of the variance for inter-income and inter-race encounters, respectively. User-level controls, including the race and income of census block groups, significantly improve the predictability of social mixing. Area control variables, region-fixed effects (Public Use Microdata Areas; PUMAs), and interaction terms also contribute to the improvement of model fit. Finally, the full model uses second-degree polynomial transformations to handle nonlinear relationships between the exogenous variables and social mixing. Overall, the full model explains over 81% and 95% of the variance observed in inter-income and inter-race encounters, respectively.

Fig. 2B shows the marginal effects of having different types of social infrastructure sites instead of baseline sites on inter-income and inter-race encounters. Marginal effects are shown for all of the types of sites (overall), for parks, community spaces, and social businesses, located in three equally sized groups of CBGs (lower, middle, upper). The mixed results (positive and negative effects) suggest that different types of social infrastructure have varying effects on inter-income and inter-race encounters. The most striking results include the strong positive associations of parks and the negative associations of community spaces with inter-income encounters. Community spaces were, on the other hand, shown to have significant positive associations with *inter*-race encounters while social businesses had negative associations with *intra*-race encounters.

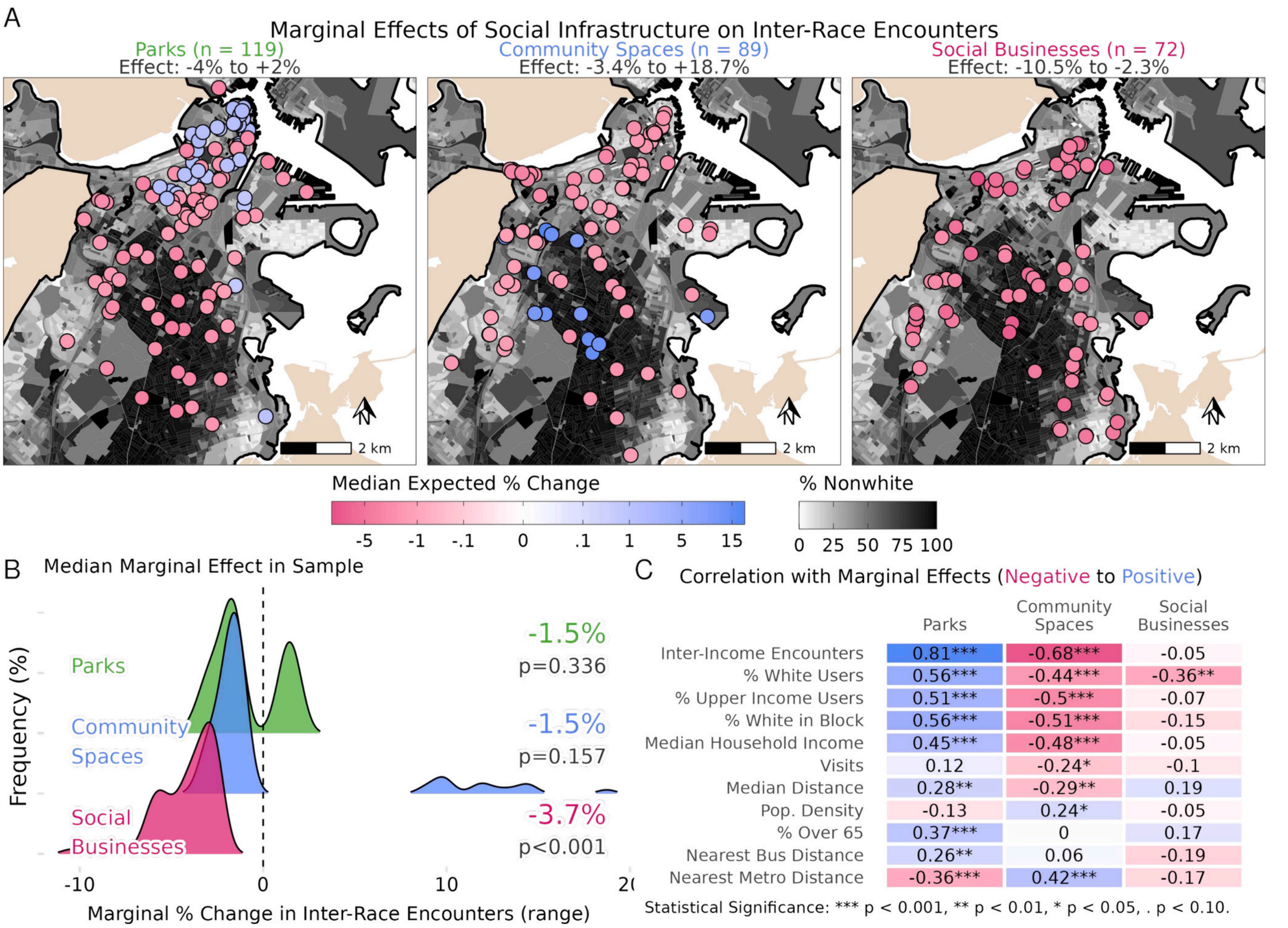

Fig. 4. Marginal effects of social infrastructure sites on inter-race encounters (compared to baseline).

Panel A maps social infrastructure sites in the Boston study area, with points colored by median expected percent change in inter-income encounters per site (compared to baseline). Median expected effects simulated using the median of 1000 simulations per site using Model 10, **Table S1**, holding all sociodemographic traits at observed values. 2020 census blocks shaded by the share of non-white residents. **Panel B**'s density curves show the full range of sites' marginal effects from the map. *P*-values show a false-positive rate (% of sites showing marginal effects on the opposite side of zero from the median effect). **Panel C** shows correlations (Pearson's r) between sites' marginal effects and sociodemographic traits.

3.2. *Marginal effects of social infrastructure sites on inter-income encounters*

How much then, did social infrastructure impact inter-group encounters in Boston? For context, in Fig. 3, we used our model to simulate the expected increase in inter-income encounters for every social infrastructure site (compared to baseline sites), conditioned on each site's observed sociodemographic traits. Our maps (**panel A**) display the locations of 119 parks (left), 89 community spaces (center), and 72 social businesses (right), shaded by the median expected percent change in inter-income encounters, while **panel B** shows the distributions of these marginal effects.

Parks had a positive association with encounters between different income groups by a median of +1.5%, relative to baseline sites ($p < 0.001$), with magnitudes ranging from +0.6% to as high as +22.9% near Roxbury. Community spaces, however, were associated with fewer inter-group encounters (−0.7%, $p < 0.001$) and instead more same-income encounters, with effects as considerable as −14.3% scattered across Boston. Finally, social businesses were positively associated with encounters between different income groups by a median of +2.5% ($p < 0.001$), with stronger effects across the city ranging from +1% to +4.3%.

What kinds of neighborhoods were positively associated with diversity? In **Panel C**, Pearson's r correlation coefficients between marginal effects and sociodemographic traits reveal that working-class, racially diverse neighborhoods tended to see stronger correlations between inter-income encounters at parks and social businesses; rates of upper-income and white residents trend negative between $r = 0.08$ and 0.29 for parks (varying significance) and − 0.45 to −0.75 for social businesses (each at $p < 0.001$). In contrast, higher inter-income encounters at community centers occurred primarily in white, wealthier neighborhoods, correlating with shares of Upper-Income users at social infrastructure sites at $r = 0.16$ (albeit lower significance).

Points, shading, distributions, tiles, correlations, and tile shading reflect the same meaning as in 3, but show simulated marginal effects on inter-race encounters for each actual social infrastructure site, conditioned on all sociodemographic variables.

3.3. *Marginal effects of social infrastructure sites on inter-race encounters over baseline sites*

In contrast, marginal effects for each actual social infrastructure site were much more varied for inter-race encounters (4) than inter-income encounters (3); parks and social businesses mapped in panel **A** show mostly negative effects on inter-race encounters (red); their benefits to same-race encounters (reflecting negative effects) were highest in predominantly non-white neighborhoods like Roxbury and Dorchester. In contrast, many community spaces in these predominantly non-white neighborhoods were positively associated with inter-race encounters (blue), while the rest had negative correlations (red).

Their distributions in panel **B** reveal that overall, most parks (66.4%), community spaces (84.3%), and social businesses (99.9%) were associated with fewer inter-race encounters and more same-race encounters than baseline sites. However, these results were most consistent and strongest for social businesses; the median social business could expect a − 3.7% decrease ($p < 0.001$) in inter-race encounters compared to baseline sites. However, panel **C** highlights that these effects are not fixed; parks' effects on inter-race encounters were strongly, significantly, positively correlated of indicators of white and wealthy neighborhoods. (Conversely, impacts of community spaces and social businesses were negatively correlated with indicators of white, wealthy neighborhoods.) This implies that wealthy, white neighborhoods were more likely to see parks that facilitated inter-race encounters, but their community spaces and social businesses tended to facilitate same-race encounters. (However, working-class, racially diverse neighborhoods tended to see community spaces and social businesses that facilitated inter-race encounters.)

4. Discussion and conclusions

In summary, this study set out to test how social infrastructure impacts the level of inter-race and inter-income encounters in communities, using 13 contiguous Boston neighborhoods as a pilot study. We hypothesized that **(H1)** social infrastructure carries value added to intergroup encounters, over ordinary sites, but that **(H2)** the type of social infrastructure matters. Indeed, we found that some types of social infrastructure in Boston were associated with more inter-group encounters: Parks were linked to more inter-income encounters, both overall ($+0.032, p = 0.016$) and at sites with more same-race encounters ($+0.062$, $p = 0.007$). Community spaces were linked to more same-income encounters (-0.025, $p = 0.046$), particularly among racially diverse individuals (upper third inter-race encounters) (-0.057, $p = 0.045$). Social businesses were linked to more same-race encounters (-0.024, $p = 0.024$), particularly given more same-income encounters ($-0.027, p = 0.068$). Community spaces were linked to more inter-race encounters given more same-income encounters ($+0.047$, $p = 0.020$).

Our findings reveal that (**H3**) inter-group mixing is strongest among certain slices of the population that are already more homogeneous in another way. But, some types of social infrastructure are linked primarily to more same-group encounters: Community spaces are linked to sizably greater same-income encounters, while social businesses are linked to more same-race encounters. Social infrastructure sites that truly bridge the whole population are quite rare, but some types (eg. parks) are better candidates for bridging than others (eg. social businesses). Also, social infrastructure's effect size varied depending on the type of encounter assessed. Social infrastructure has the largest effects on inter-*income* encounters (up to 6%), but smaller effects on inter-*race* encounters (up to 4%, with most at 1%), which is so greatly shaped by the race and income of users and the surrounding area.

Our results make several contributions to the literature on social infrastructure and inter-group social ties, providing empirical evidence for the first time on two major questions. First, our findings identified what kinds of social infrastructure, given what kind of demographic situations, alleviates income-segregation. Our results suggest that parks and social businesses alleviate income segregation and encourage inter-income mixing particularly in working class neighborhoods and racially diverse neighborhoods (as show in 3 panel C). Community spaces do not alleviate income segregation at large, but their negative impacts decline in white, wealthy neighborhoods. Further, parks and social businesses pose a potential tool for encouraging encounters, but these are primarily same-race encounters.

Second, our findings help clarify that the benefits of social infrastructure vary substantially, more so that scholars may think. For example, if a site bridges inter-income groups (but not inter-race groups), or if a site bridges inter-race groups (but not inter-income groups), these patchwork gains still do help bridge residents across the city. These increased diverse encounters from social infrastructure matter, because the more diverse ties residents have, the easier it is for them to navigate and recover from shocks. Cities face increasingly frequent shocks in response to climate change-induced hazards, the pandemic, and refugee crises; when these shocks strike, residents with greater trust and reciprocity across group lines are more likely to share resources, come to each others' aid, and recover faster (Aldrich, 2012; Lee et al., 2022). Fortunately, our research indicates that some types of social infrastructure, especially parks and social businesses, may facilitate these inter-income and inter-race encounters better than other ordinary settings (in this study, better than grocery stores, supermarkets, or post offices). Similarly, some types of social infrastructure are designed to serve the local community more than others (e.g., community spaces like libraries and neighborhood centers); consequently, it is not surprising that these spaces tend to promote more homophilous interactions, likely connecting just demographics within that neighborhood. In this way, perhaps we should not require social infrastructure to be a panacea for all types of social ties by default; certain social infrastructure types can still make meaningful impacts even when reaching just subsets of the population.

The results in this study have certain limitations due to the nature of the mobility data. Since mobility data only provides the visitation patterns, we are not able to understand the intent of the visit to these places, or how engaging they were with others during the visit. Using more microscopic data sources that can capture interpersonal transactions and dynamics, including survey data and interview data, could enhance our analysis. Moreover, in this study, we focused on understanding the encounters before the pandemic. A recent study has revealed that the diversity of encounters in cities has substantially decreased during the pandemic due to long-term behavioral changes (Yabe et al., 2023). A natural extension of this study would be to understand the role of social infrastructure places during the different stages of the pandemic. Mobile phone location data collected through smartphone apps, is known to contain various types of biases, such as sociodemographic, spatial, temporal (data collection frequency) differences in representativeness (Yabe et al., 2024). In this study we corrected for such biases through post-stratification of individual users inversely proportional to their home census block group's sampling rate, and the study could benefit from further work on addressing this issue from multiple dimensions. In this study, we quantified the marginal benefits of having a social infrastructure site over a baseline site using econometric modeling approaches. A potential future research direction could be to find locations that were renovated into public spaces and to quantify the causal impact of the investment across time.

Code availability

The analysis was conducted in Python and R. Code to reproduce the main results in the figures from the aggregated data will be made publicly available on GitHub.

Funding

T.Y. acknowledges support through NSF grant No. 2343646. D.A. acknowledges support of NSF grant 2228533 and the Rasmussen Foundation. E.M. acknowledges support through NSF grants No. 2218748 and 2420945.

CRediT authorship contribution statement

Takahiro Yabe: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Timothy Fraser:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Daniel P. Aldrich:** Writing – review & editing, Writing – original draft, Supervision, Resources, Investigation, Conceptualization. **Esteban Moro:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors have no competing interests to declare.

Data availability

The data that support the findings of this study are available from Spectus through their Social Impact program, but restrictions apply to the availability of these data, which were used under the license for the current study and are therefore not publicly available. Information about how to request access to the data and its conditions and limitations can be found in https://spectus.ai/social-impact/.

Acknowledgements

We would like to thank Spectus who kindly provided us with the mobility dataset for this research through their Data for Good program.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compenvurbsys.2024.102173.

References

Aldrich, D. P. (2012). *Building resilience: Social Capital in Post-Disaster Recovery*. University of Chicago Press.

Aldrich, D. P., & Meyer, M. A. (2015). Social capital and community resilience. *American Behavioral Scientist, 59*(2), 254–269. https://doi.org/10.1177/0002764214550299

Aleta, A., Martin-Corral, D., Piontti, A. P. Y., Ajelli, M., Litvinova, M., Chinazzi, M., … Moreno, Y. (2020). Modelling the impact of testing, contact tracing and household quarantine on second waves of covid-19. *Nature Human Behaviour, 4*(9), 964–971.

Blondel, V. D., Decuyper, A., & Krings, G. (2015). A survey of results on mobile phone datasets analysis. *EPJ Data Science, 4*(1), 10.

Blumenstock, J. E. (2012). Inferring patterns of internal migration from mobile phone call records: Evidence from Rwanda. *Information Technology for Development, 18*(2), 107–125.

Chetty, R., Hendren, N., Jones, M. R., & Porter, S. R. (2019). Race and economic opportunity in the United States: An intergenerational perspective*. *The Quarterly Journal of Economics, 135*(2), 711–783. https://doi.org/10.1093/qje/qjz042

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94*, S95–S120.

Cribari-Neto, F., & Zeileis, A. (2010). Beta regression in R. *Journal of Statistical Software, 34*(2), 1–24. https://doi.org/10.18637/jss.v034.i02

Fraser, T. (2022). Simulate: Tidy simulation for statistical models, r package version 0.0.0.9000. URL https://github.com/timothyfraser/simulate.

Fraser, T., Cherdchaiyapong, N., Tekle, W., Thomas, E., Zayas, J., Tan, C., & Aldrich, D. P. (2022). *Trust but verify: Validating new measures for mapping social infrastructure in cities, Tech. rep*. Social Science Research Network. https://doi.org/10.2139/ssrn.4065162

Fraser, T., & Naquin, N. (2022). Better together? The role of social capital in urban social vulnerability. *Habitat International, 124*, Article 102561.

Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A.-L. (2008). Understanding individual human mobility patterns. *Nature, 453*(7196), 779.

Greenberg, E., & Parks, R. P. (1997). A predictive approach to model selection and multicollinearity. *Journal of Applied Econometrics, 12*(1), 67–75.

Hanibuchi, T., Kondo, K., Nakaya, T., Shirai, K., Hirai, H., & Kawachi, I. (2012). Does walkable mean sociable? Neighborhood determinants of social capital among older adults in Japan. *Health & Place, 18*(2), 229–239.

Hwang, J., & Ding, L. (2020). Unequal displacement: Gentrification, racial stratification, and residential destinations in Philadelphia. *American Journal of Sociology, 126*(2), 354–406. https://doi.org/10.1086/711015. URL https://doi.org/10.1086/711015.

Imai, K., King, G., & Lau, O. (2008). Toward a common framework for statistical analysis and development. *Journal of Computational Graphics and Statistics, 17*(4), 892–913. URL https://gking.harvard.edu/files/abs/z-abs.shtml.

Jacobs, J. (1993). *The death and life of great American cities*. Vintage Books.

Johnson, C. A. (2010). Do public libraries contribute to social capital?: A preliminary investigation into the relationship. *Library & Information Science Research, 32*(2), 147–155.

Kawachi, I., Subramanian, S. V., & Kim, D. (2008). *Social capital and health*. Springer.

Kim, J., & Kaplan, R. (2004). Physical and psychological factors in sense of community: New urbanist kentlands and nearby orchard village. *Environment and Behavior, 36*(3), 313–340.

King, G., Tomz, M., & Wittenberg, J. (2000). Making the most of statistical analyses: Improving interpretation and presentation. *American Journal of Political Science*, 347–361.

Klinenberg, E. (2018). *Palaces for the people: How social infrastructure can help fight inequality, polarization, and the decline of civic life*. Crown.

Latham, A., & Layton, J. (2019). Social infrastructure and the public life of cities: Studying urban sociality and public spaces. *Geography Compass, 13*(7), Article e12444.

Lee, J., Aldrich, D. P., Kiyota, E., Yasuhiro, T., & Sawada, Y. (2022). Social capital building interventions and self-reported post-disaster recovery in ofunato, Japan. *Scientific Reports, 12*(1), 1–10.

Leyden, K. M. (2003). Social capital and the built environment: The importance of walkable neighborhoods. *American Journal of Public Health, 93*(9), 1546–1551.

Lund, H. (2003). Testing the claims of new urbanism: Local access, pedestrian travel, and neighboring behaviors. *Journal of the American Planning Association, 69*(4), 414–429.

Moro, E., Calacci, D., Dong, X., & Pentland, A. (2021). Mobility patterns are associated with experienced income segregation in large us cities. *Nature Communications, 12*(1), 1–10.

Podobnik, B. (2011). Assessing the social and environmental achievements of new urbanism: Evidence from Portland, Oregon. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability, 4*(2), 105–126.

Putnam, R. D. (2000). *Bowling Alone. The Collapse and Revival of American Community, Simon and Schuster, New York*.

Recht, H. (2020). Censusapi: Retrieve data from the census APIs, r package version 0.7.1. URL https://github.com/hrecht/censusapi.

Rogers, S. H., Halstead, J. M., Gardner, K. H., & Carlson, C. H. (2011). Examining walkability and social capital as indicators of quality of life at the municipal and neighborhood scales. *Applied Research in Quality of Life, 6*(2), 201–213.

Small, M. L. (2017). *Someone to talk to*. Oxford University Press.

Szreter, S., & Woolcock, M. (2004). Health by association? Social capital, social theory, and the political economy of public health. *International Journal of Epidemiology, 33*(4), 650–667. https://doi.org/10.1093/ije/dyh013. URL https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyh013.

E. Villalonga-Olives, T. Wind, I. Kawachi, Social capital interventions in public health: A systematic review, Social Science & Medicine 212 (2018) 203–218. doi:https://doi.org/https://doi.org/10.1016/j.socscimed.2018.07.022.

Walker, K., & Herman, M. (2021). Tidycensus: Load US census boundary and attribute data as 'tidyverse' and 'sf'-ready data frames, r package version 1.1. https://walker-data.com/tidycensus/.

Wood, L., Frank, L. D., & Giles-Corti, B. (2010). Sense of community and its relationship with walking and neighborhood design. *Social Science & Medicine, 70*(9), 1381–1390.

Yabe, T., Bueno, B. G. B., Dong, X., Pentland, A., & Moro, E. (2023). Behavioral changes during the covid-19 pandemic decreased income diversity of urban encounters. *Nature Communications, 14*(1), 2310.

Yabe, T., Jones, N. K., Rao, P. S. C., Gonzalez, M. C., & Ukkusuri, S. V. (2022). Mobile phone location data for disasters: A review from natural hazards and epidemics. *Computers, Environment and Urban Systems, 94*, Article 101777.

Yabe, T., Luca, M., Tsubouchi, K., Lepri, B., Gonzalez, M. C., & Moro, E. (2024). Enhancing human mobility research with open and standardized datasets. *Nature Computational Science*, 1–4.