# Diversity beyond density: Experienced social mixing of urban streets

Zhuangyuan Fan [a,b,*], Tianyu Su [c,d], Maoran Sun [a,e], Ariel Noyman[f], Fan Zhang [g], Alex 'Sandy' Pentland[f] and Esteban Moro [f,h,i,*]

[a]Senseable City Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[b]Department of Geography, The University of Hong Kong, Pok Fu Lam, Hong Kong SAR
[c]Graduate School of Design, Harvard University, Cambridge, MA 02138, USA
[d]Institute of Quantitative Social Science, Harvard University, Cambridge, MA 02138, USA
[e]Department of Architecture, The University of Cambridge, Cambridge, UK
[f]Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[g]Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong, Hong Kong SAR
[h]Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
[i]Grupo Interdisciplinar de Sistemas Complejos (GISC), Departamento de Matemáticas, Universidad Carlos III de Madrid, Leganés 28911, Spain
*To whom correspondence should be addressed: Email: yuanzf@mit.edu (Z.F.), emoro@mit.edu (E.M.)
**Edited By:** Taylor Jaworski

## Abstract

Urban density, in the form of residents' and visitors' concentration, is long considered to foster diverse exchanges of interpersonal knowledge and skills, which are intrinsic to sustainable human settlements. However, with current urban studies primarily devoted to city- and district-level analyses, we cannot unveil the elemental connection between urban density and diversity. Here we use an anonymized and privacy-enhanced mobile dataset of 0.5 million opted-in users from three metropolitan areas in the United States to show that at the scale of urban streets, density is not the only path to diversity. We represent the diversity of each street with the experienced social mixing (ESM), which describes the chances of people meeting diverse income groups throughout their daily experience. We conduct multiple experiments and show that the concentration of visitors only explains 26% of street-level ESM. However, adjacent amenities, residential diversity, and income level account for 44% of the ESM. Moreover, using longitudinal business data, we show that streets with an increased number of food businesses have seen an increased ESM from 2016 to 2018. Lastly, although streets with more visitors are more likely to have crime, diverse streets tend to have fewer crimes. These findings suggest that cities can leverage many tools beyond density to curate a diverse and safe street experience for people.

**Keywords:** social mixing, income segregation, built environment, computer vision, human mobility

## Significance

The concern about rising inequality has intensified the interest in how cities can build a socially mixed environment. With the current progress in urban studies, we still cannot unveil the elemental connection between urban environment, density, and diversity. This study leverages a mobile dataset that contains 0.5 million users to measure income mixing at the street level. We show that the concentration of visitors only explains up to 26% of street mixing, while the adjacent amenities, residential mixing, and income level would account for more than 44% of social diversity. Our results highlight the importance of street-level factors that influence social mixing and diversity and can be useful for urban planners and policymakers to create more socially mixed environments.

## Introduction

Diversity is intrinsic to a sustainable, resilient, and inclusive city (1, 2). Within many forms of diversity, the diverse collection of people, socially and economically, is one of the crucial preconditions of economic urban vitality (3) and creativity (4–6). On the contrary, the segregation of people was shown to impact children's economic outcomes (7), widening the digital gap and hindering access to public health services (8). Correspondingly,

researchers across the fields of economics, sociology, urban planning, and mobility have devoted themselves to explaining the level of social mixing and segregation across space and time.

Although most studies of vitality and mixing in our cities has been done using static, residential-only, and sometimes outdated census data, recent studies of human mobility draw our attention to the activity space in cities beyond where people live. It is clear that people do not only stay where they live but also work, travel, and relax in places other than their homes. Therefore, people

living in less diverse neighborhoods may still have chances to encounter people with different demographics and knowledge during their daily life. Recent studies have measured how well people with different backgrounds are mixed during their daily travel activities (9–11), online communications, and purchase activities (12). It is shown that the likelihood of people meeting diverse others is related to an individual's demographic characteristics, lifestyle, and travel habits (10). However, beyond an individual's behavior choices, what remains unanswered is how a city as a system could build an environment that cultivates social mixing in the long run.

We yet still face one gap in addressing this question. While most studies on social mixing, segregation, and diversity were conducted at the city or district level, planning in practice concerns more with the fine spatial resolution of public space in cities. Building on the work that creates an activity-based measure of diversity and segregation, we address this study gap by focusing on the space of street sidewalks. Theoretically, street sidewalks are critical urban open spaces advocated by sociologist William H. (Holly) Whyte (13), journalist Jane Jacobs (1), architect Jan Gehl (14), and New Urbanism scholars (15). Practically, the United Nations Sustainable Development Goals' target 11 emphasizes the vital role of urban public spaces in social and economic life.

To understand the vital elements leading to socially mixed street experience, we first create a measure of experienced social mixing (ESM) that estimates the income aspect of mixing in cities using a large collection of micro-scale mobility data across 40 counties and 3 metropolitan areas in the United States. The ESM measures the evenness of time spent at each street segment by people from different income groups. This measure describes the experience of diversity when people visit a street segment that is not immediate in their living neighborhood. It is different from the "vitality" described by Jane Jacobs (2) as we do not observe the various types of activities people conduct together. Instead, we capture the likelihood of people from different income groups co-presence at a street segment in their daily life.

Many urban theories and practices have argued the importance of density, or the concentration of people, in leading towards more urban vibrancy (2, 3, 16, 17), and thus favoring a more socially mixed urban environment (4, 18, 19). If we view the desired outcome as the diverse admix of human knowledge, abilities, preferences, interaction, and so forth, using density as the only tool has its limitations—city blocks with a high density of office buildings could still only see people with similar income levels and skill sets. With the relationship between density and diversity inevitably being non-linear, we should further unpack what other tools cities could leverage to curate a socially mixed urban environment.

Therefore, here we examine what factors beyond density could further contribute to social mixing. Here we represent density with the total number of visitors to each street segment. Two main sets of factors connected with human mobility and social interactions are examined in this study. We first discuss the importance of socioeconomic factors, including income, residential density, and residential mixing. These variables are learned from the human mobility and segregation literature that people tend to visit places at a given income segregation level (10), and residential segregation correlates with experienced segregation at a city scale (9). The second set of factors describes venues along the streets and how safe the street looks. These variables resonate with the city planning literature that advocates the mixed-use development (20–22) and street environment safety (2, 23, 24).

We present three main results. First, conditioning on density, ESM can still be explained by adjacent neighborhoods' residential mixing, income level, and venues along the street. Density, in our measure, the number of visitors visiting a street segment at any specified time, only explains around 26% of the model estimation, while residential mixing, income level, and venues contribute to 44% of the model estimation.

Second, ESM measured at different hours of a day is closely related to different types of venues along a street segment, highlighting the importance of a mixed-use environment. Among the venues, food-related businesses exert the highest contribution to explaining the ESM at different times of the day. Meanwhile, when controlling for the total number of visitors, the streets with more coffee and tea venues can attract more diverse groups of people.

Moreover, we found that the street segments with an increase in food-related business from 2016 to 2018 are likely to see an increase in the ESM. This longitudinal effect holds conditioning on the increase of total visitors.

The well-being of urban dwellers is a multi-dimensional concept that goes beyond diversity and economic status and involves health, crimes, and other aspects of life. Beyond the daily ESM, a body of recent literature indicates that mobility patterns also predict crimes (25–27)—where residents visit is also a source of neighborhood (dis)advantage (28). To further understand the effect of the ESM, the last part of our study analyzes the relationship between visitor volumes, ESM, and different types of crime incidents around each street segment. We show that although denser cities attract more crime incidents, conditioning the visitor volumes, street-level ESM has a negative association with crime count.

This study highlights the importance of a high-resolution measure of social mixing as a spatial-temporal dynamic urban phenomenon—streets adjacent to each other could present dramatically different levels of ESM at different times of the day. Furthermore, we illustrate how cities could leverage the open space of the street sidewalks to increase the chance for different people to meet each other and thus mitigate the existing downfall of residential segregation. Lastly, our result also shows that diverse visitor experience does not always go in parallel with high volumes of crime incidents. Large cities can leverage many policy tools beyond density to curate a diverse and safe street experience for people.

## Results
### Street-level ESM

We create a measure of ESM for three large metropolitan areas of Boston, New York, and Philadelphia, which involves more than 40 counties across 5 states (MA, NY, DL, NJ, PA). The privacy-enhanced mobility data are provided by Cuebiq, which includes 3-month long records across 2 years of anonymized device-level location pings for 0.5 million users who opted into data sharing for research purposes under a General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA)-compliant framework.

To construct the ESM, we pre-process the data to identify each device's home census block group (CBG) and stay locations using the same method in a previous work by Moro et al. (10). We first associate each device from the mobility dataset with an approximate socioeconomic status by their inferred home CBG. Each individual's home CBG is obtained from their most commonly visited location between 10 PM and 6 AM (see Methods section). Then, all individuals are grouped into four quantiles of income groups
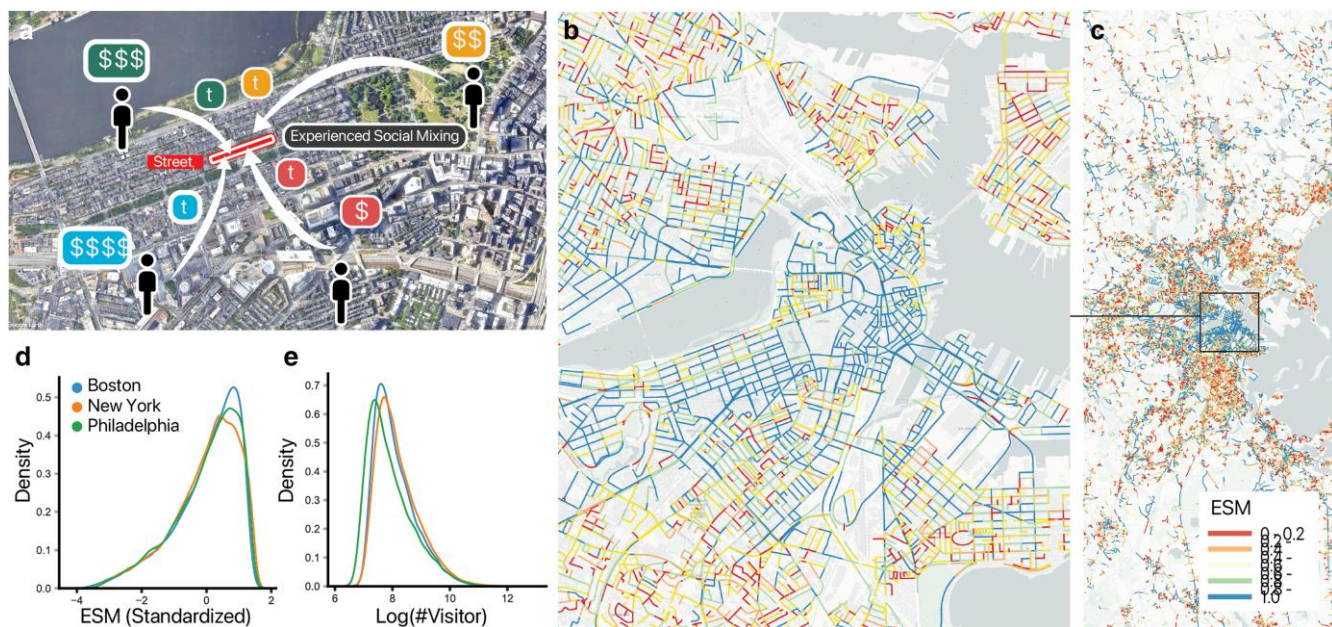
**Fig. 1.** Measure the ESM. a) Street-level ESM is calculated based on the time spent by each income group at each street segment. Basemap is exported from Google Earth. b) Distribution of daily ESM of Boston. Only street segments with at least 1 POI within the 100-m buffer area are visualized on the map. c) Regional map visualizing the ESM around Boston metropolitan area. d) Distribution of the ESM by metropolitan area. e) Distribution of log-transformed visitor counts during the study period by metropolitan area. Basemaps are from QGIS HCMGIS plugin.

according to their home CBG's median household income's relation to the metropolitan area distribution of median household income (see Methods section). We then extract visits an individual made to a given street segment for at least 5 minutes but a maximum of 2 hours. This is to prioritize sidewalk activities that have the potential for meaningful interaction among pedestrians. Activities such as visiting cafes, restaurants, and parks, or simply resting along the streets are emphasized. Other activities such as working in an office building or watching movies, which usually take a long time but offer little chances for people to meet each other, are dropped. The post-stratification process reduces sample bias regarding population and income level (see Supplementary Note 2 and Figs. S2 and S3a).

We compute the street-level ESM with the proportion of total time spent at that street by each income quartile (see Fig. 1a). The ESM is defined as the Shannon entropy [29] of each income group's activity time spent at a given street segment (see Methods section). ESM quantifies income mixing from 0 to 1. A street segment that is fully mixed ($ESM_s = 1$) when the total time across all individuals spent at the street segment is split evenly among the four income quartiles, while a street segment with $ESM_s = 0$ indicates that the street segment is only visited by one income group. Fig. 2b shows several examples. The bar chart of each street represents the accumulated time spent by each income group at the sample street. The dashed line demonstrates when the street would be fully mixed ($ESM_s = 1$). For example, 63 Cambridge Street in Boston has a lower ESM (0.767) in comparison to 27 Tremont Street (0.975), as the time spent by each income group along 27 Tremont Street is more evenly distributed. We note that there are other choices for mixing and segregation metrics. We also examine the robustness of our measure of social mixing against other metrics (see Supplementary Note 3 and Fig. S3b).

Street-level ESM presents spatial heterogeneity in each city (Fig. 1c). However, it also has a very fine-grained spatial resolution. We illustrate this observation in Fig. 2a and b. Fig. 2a presents the

proportion of time spent by each income group to each street segment in the Boston downtown area. Each dot represents 0.5% of time spent by each income group. Fig. 2b demonstrates the street view samples and their associated distribution of time spent by each income group. Even two adjacent streets with the same intersection could present drastically different levels of social mixing. This finding shows that understanding vitality, density, or mixing in our cities at larger scales (e.g. Census Tracts or districts [30]) misses the fine-grained structure of how people interact and encounter in our cities and the relationship of the diversity of encounters and urban environment.

## Explain the street-level ESM

To understand the relationship between the streets and ESM, we model the ESM of each street using a spatial autoregressive model with generalized spatial two-stage least-square (G2SLS) estimates. The explanatory variables include the street segment's length, the segment's distance from the closest metropolitan center, the street type (one-way vs. two-way), street sinuosity, a series of residential factors that describe the adjacent neighborhood in which the street segment is located, the number of venues (such as restaurants, cafes, grocery stores, schools, etc.) located along the street, and how safe a street segment look. The residential factors include population density, median household income, and residential income mixing. In addition, considering the geographical differences of all samples included in the study, we also add county-level fixed effects. To compute the residential factors, we obtain a collection of CBGs that fall within a street segment's 800-m buffer and use the median household income and population size of each CBG accordingly (see Methods section and Supplementary Note 4.4). The residential income mixing mirrors the calculation of ESM. It measures the level of income mixing of the four income groups, where 0 indicates that all residents within the 800-m buffer belong to the same income group,
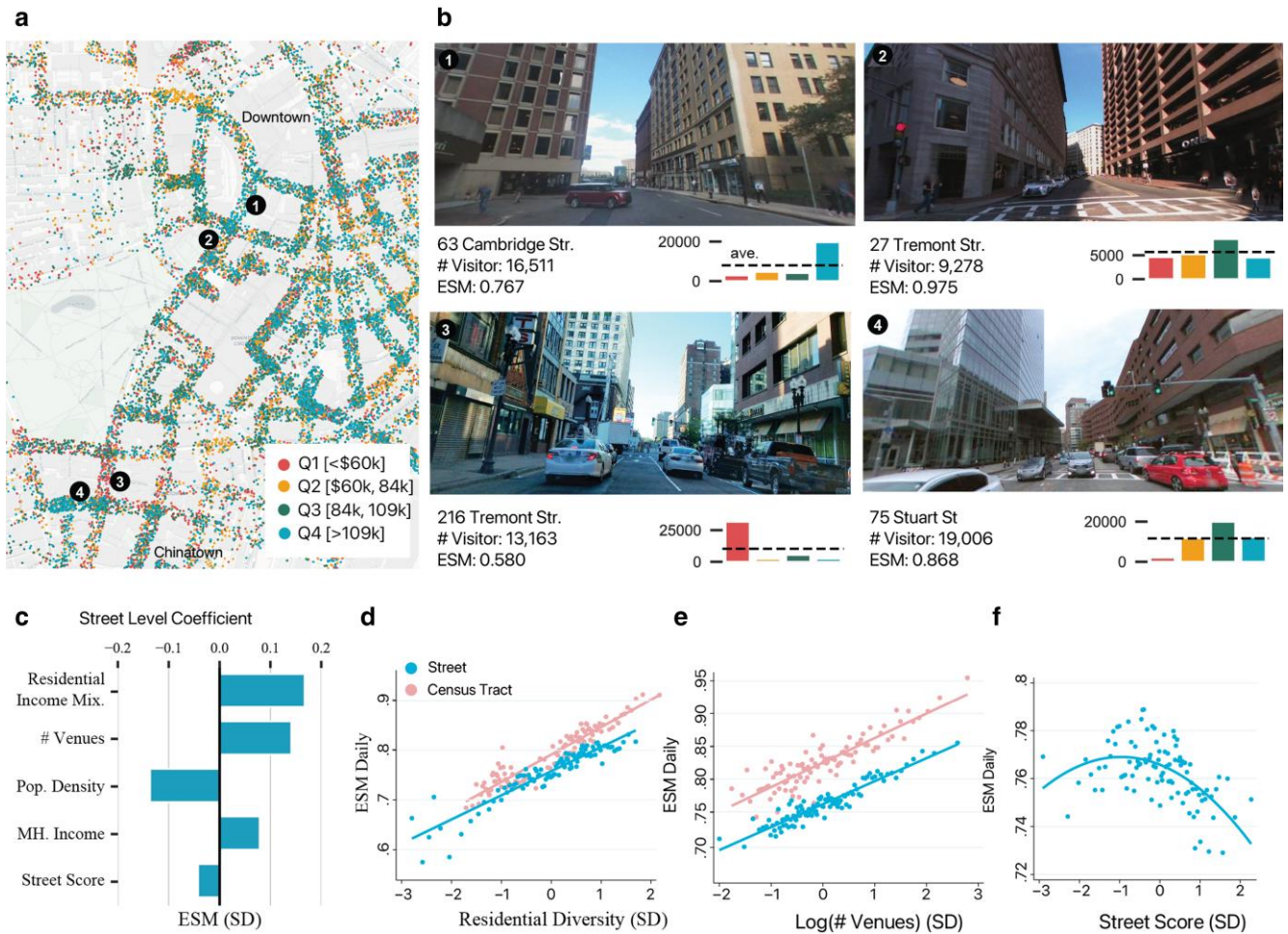
**Fig. 2.** Explain the ESM. a) Street-level ESM representation. Each dot represents 0.5% of time spent at a street segment by a group of visitors. b) Street view examples of streets that are adjacent to each other yet present different ESM levels. Street view images were downloaded from Mapillary. Each barplot presents the hour of visits from each income group. The total number of unique visitors is shown accordingly. c) Coefficients of all variables estimating street-level social mixing (all shown variables have *P*- value < 0.005) based on Eq. 5. d) Compare the effect of residential mixing on ESM at street and CT levels. Binned scatter plot with 100 quantiles. e) Compare the effect of the number of POI on ESM at the street and CT levels. f) A non-linear relationship between the Street Score and the street-level ESM (see the full table in Supplementary Table S1).

whereas 1 implies a fully mixed street residential neighborhood (see Methods section).

The venues on each street are obtained from Foursquares' point of interest (POI) API. A collection of 0.1 million verified venues across the study areas are used. Lastly, given that the sense of safety is one of the significant concerns determining the street activity (2), we measure how safe each street looks through the *Street Score*. The *Street Score* of each street segment is predicted from the Google Street View (GSV) images taken along the street segment. The *Street Score* model was adapted from Ref. (31), which is a convolutional neural network trained with the data from Place Pulse (32). We collected 1.6 million street view images through Google API across the study areas, and images taken from winter were dropped to avoid seasonal inconsistency. A similar method was used by Naik et al. (24) in describing the inequality of urban safety perception.

Fig. 2c summarizes the regression coefficients of the above-mentioned linear regression. Besides the geographical fixed effect, street segments close to a higher level of residential mixing and more venues tend to be more socially mixed. One standard deviation (SD) increase in residential income mixing is associated with a 0.17 SD increase in ESM. We test if this relationship holds

at the different spatial units by repeating the experiment at Census Tract (CT) level. Fig. 2d and e plots the results in parallel. We find that the CTs with more mixed residential composition and venues also tend to be more mixed.

We also find that population density and *Street Score* are negatively associated with the street-level ESM. One SD increase in population density is associated with a 0.14 SD decrease in ESM. The former indicates that neighborhoods with more residents do not guarantee more chances of cross-group mixing during their daily activities. To further unpack the negative association between the *Street Score* and the street-level ESM, we included a quadratic term of the safety score in the same model and identify a non-linear relationship between the *Street Score* and the ESM (see Fig. 2f). This could reflect a number of forces, including gentrification and fear of crime in cities. People tend to avoid places that look very unsafe (33), but in the meantime, streets with luxury settings, well-planted groves of trees, and fine furniture also indicate a sense of gentrification do not welcome social mixing (34). It is worth noting that the effect of *Street Score* diminishes at CT-level study, implying that the perception of street view matters more at a scale of the street segment rather than in larger spatial unit (see Supplementary Table S1).
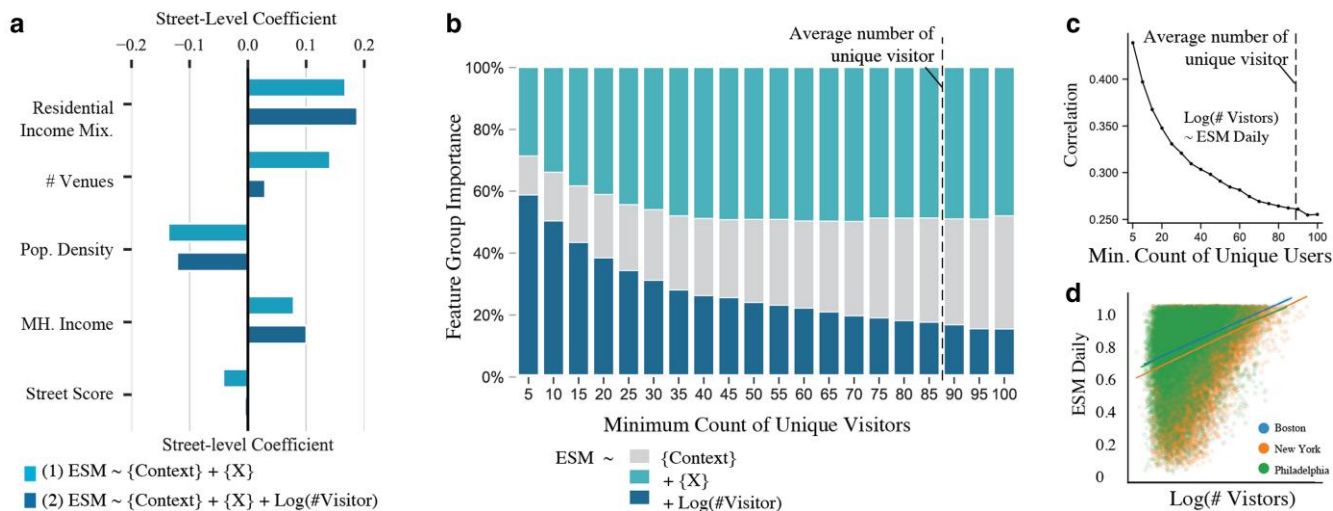
**Fig. 3.** How much does the concentration of visitors explain social mixing? a) Coefficients of all variables estimating the ESM (standardized) with and without the number of visitors controlled (spatial network spillover effect of ESM considered). *Context* variables include the geographical fixed effects, segment's distance from the closest metropolitan center, the street type, and street sinuosity; *X* include the residential income mixing, income level, population density, the number of venues, and *Street Score*. b) Summary of importance of feature groups in the linear model for social mixing by sequentially dropping street samples visited by fewer count of unique users (Eq. 4). c) Correlation of log-transformed number of visitors and daily ESM by sequentially dropping street samples visited by fewer count of unique users. d) Relationships between the log-transformed visitor counts and daily ESM by different metropolitan areas (see Supplementary Fig. S5 for results comparison).

## ESM and density

As many current urban theories indicate, urban density is one of the important instruments that foster diversity (2, 3, 16, 17). A street segment with more visitors could naturally have a higher chance of being more socially mixed. Could it be that the factors we discussed above are associated with the ESM through the channel of density?

As we can see in Fig. 3d, the log-transformed total number of visitors visiting each street has a correlation with the street-level ESM in different cities. We find this correlation dropped as we sequentially excluded street segments with too few unique visitors to avoid the small sample bias (Fig. 3c). This suggests that the ESM is not solely explained by density, albeit they are correlated. Fig. 2b also shows specific street segment examples and implies that streets with fewer total visitors can still be more mixed than others.

To further quantify the factors explaining ESM beyond density, we repeat the regression model in the previous section by including the log-transformed count of visitors as a variable. Fig. 3b shows the importance of variable groups in predicting street-level ESM. By excluding street segments with too few unique visitors (fewer than 20), the count of visitors accounts for around 26% of the variance in street-level ESM. Apart from the geographical fixed effect and street segment length, the residential income mixing, income level, population density, and venues account for around 44% of the model variance.

Again, Fig. 3a plots the comparison of two models (see Supplementary Table S1 for the full result). We show that by including the log-transformed total visitor count as a variable, the effects of residential income mixing, income level, and population density hold. The effect of the total count of venues and the *Street Score* dropped drastically. Specifically, before including the total number of visitors, a 1 SD increase in the number of venues is associated with a 0.14 SD increase in ESM. After including the number of visitors, a 1 SD increase in the number of venues is associated with a 0.03 SD increase in ESM. These observations lead to two insights: one is that the streets with more venues and look less safe tend to

have more visitors and, therefore, are more socially mixed. The other lies in the fact conditioning the ability to attract more visitors, street segments within areas that have more mixed residential environments and higher income levels are still more socially mixed. One might wonder if the mixed residential environment contributes to the ESM directly as a street segment's neighborhood residents would visit the street segment very often. We calculate the average travel distance from any given street to its visitors' home CBG's geometry center. We found that more than 95% of the street segments' visitors live more than 800 m away from the street. This result confirms that the street segment with a diverse residential environment attracts visitors from different income groups even though they do not live nearby.

## Temporal variation of ESM

Street life has a unique temporal characteristic (2, 35). To better understand the temporal variation of ESM throughout the day, we reconstruct the model by estimating the ESM at four selected time periods of a day. Fig. 4a shows the average street-level ESM for each metropolitan area throughout the day. We observe that, on average, the ESM is highest around noon and lowest in the morning, reflecting the daily activities in cities. In addition, we group venues by their categories to test how the category of venues might predict ESM dynamically (see Supplementary Table S5 for venue summary by types). Similarly, we compare the model results by excluding and including the count of visitors at different times correspondingly (Fig. 4).

Fig. 4b shows that streets with more venues such as food, coffee, and tea are more mixed throughout the day, while the streets with more shopping and entertainment (bars and clubs) venues become more mixed in the late afternoon. Streets with more health-related venues tend to be more mixed before 6 PM, and a similar trend is also seen for streets with more work-related venues. We interpret these results to be associated with both the schedule of the business and people's mobility patterns. While people mostly visit hospitals and clinics during the day, streets with these venues tend to be more mixed during their normal
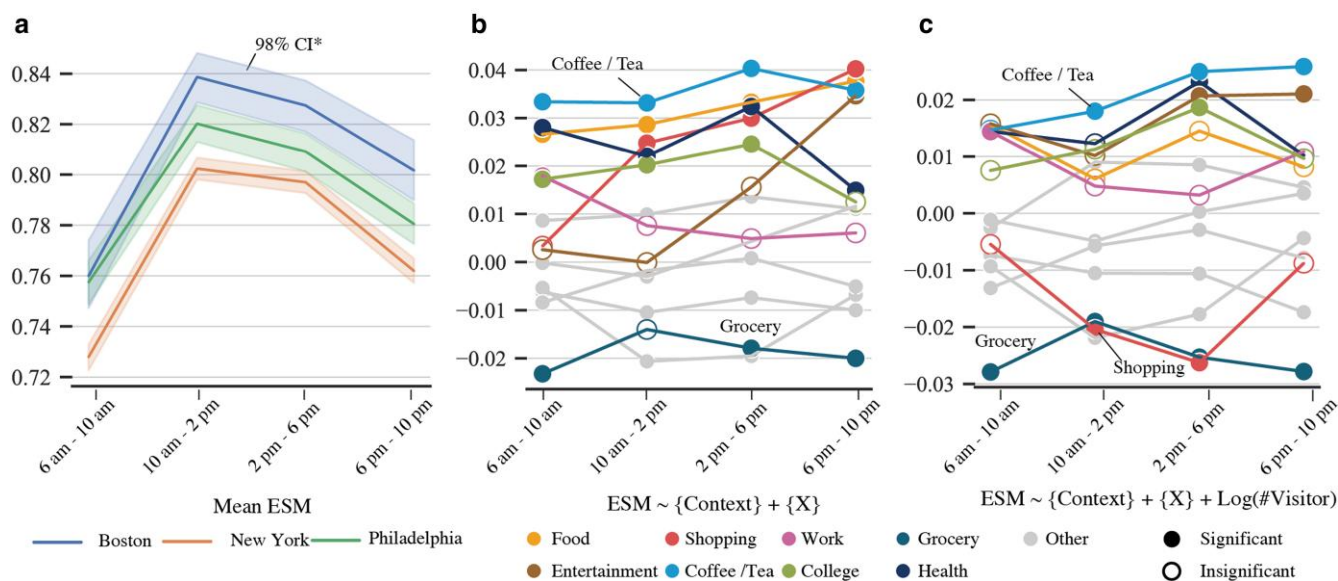
**Fig. 4.** Time-variant ESM and street venues. a) Mean time-variant ESM for each metropolitan area. Note that we only present the streets with at least 20 unique visitors at each time period of time. CI* stands for the confidence interval. Ninety-eight percent of the confidence interval is shaded. b) Time-variant effects of the number of venues by their categories. c) Time-variant effects of the number of venues by their categories conditioning on the visitor count (spatial network spillover effect of ESM considered in b and c. Only coefficients with P value smaller than 0.05 are shown as significant. Full table results and the coefficient matrix are shown in Supplementary Table S2 and Fig. S6.

operating hours. Since people will be more likely to visit bars and go shopping after work, streets with these venues only start to see mixed groups of people later in the afternoon. We also find that streets with more grocery stores tend to have consistently lower levels of mixing throughout the day. This is likely an effect of segregated residential neighborhoods, where people tend to go to grocery stores closer to where they live.

In parallel, we also add the number of visitors to each street segment to test if the impact of density would change the model results. Fig. 4c shows that conditioning on the count of visitors and the number of coffee, tea, food, entertainment, and health care venues are still contributing to more mixed streets throughout the day, while streets with more grocery stores are still less diverse. We understand that healthcare facilities are naturally more integrated, given their unique service role in the city. However, the effect of coffee, tea, food, entertainment, and grocery stores reveals that even on streets with a similar amount of visitors, their level of social mixing can still vary considering the different functions it provide. After controlling the density, we also found that the effect of shopping venues dropped and even reversed. We interpret this observation as that shopping venues are more successful in bringing in a high volume of people, yet the people group attracted to these areas might not be as diverse.

## Changes of ESM from 2016 to 2018

The cross-sectional study above highlights the food-related venues in predicting street-level ESM. We further design an experiment to test if the relationship holds longitudinally. Here, we leverage a crowd-sourced dataset, Boston's Hidden Restaurant,[a] contributed by local communities from the Boston region to test if the open and close of food-related businesses from 2016 to 2018 cast any impact on the changes of street-level ESM in the corresponding time, controlling for the changes of residential features (summary stats included in Supplementary Table S3). Understanding that streets with a very high ESM in 2016 would

have less room to improve than streets that were less mixed, we control for the ESM at 2016 for all models. In addition, the model also includes the same social and geographical context features in 2016 to account for the potential trend differences (see the full result in Supplementary Table S4).

Table 1 illustrates the results. Column 1 indicates that among all residential variables, the change in the proportion of residents with at least a bachelor's degree is the only feature that contributes to the change in ESM. Specifically, 1 SD increase in population with at least a bachelor's degree is associated with a 1.2% SD increase in ESM. With all other features controlled, we found little relationship between the changes in residential income diversity and changes of ESM. This is partly because the residential income diversity only changes very subtly between the 2 years.

Columns 2 and 3 indicate that the absolute increase in the number of food businesses positively correlates with the change in ESM. One SD more food-related business is associated with a 0.3% SD increase of ESM. The coefficient of change of education level still holds by including the change in the food business. Column 4 includes an interaction term to test the marginal effect of the food business, considering that streets with different original ESM in 2016 might respond to the changes differently. We show that the interaction term of ESM 2016×Δ food businesses is negatively associated with the change of ESM. It implies that with a similar increase of food businesses, the streets with a lower ESM in 2016 tend to see more increase of ESM (Fig. 4a). Fig. 4b shows examples of streets with changes in the food business and different levels of ESM in 2016. Lastly, we further include the log-transformed changes of visitors to each street segment between 2016 and 2018 in column 5. As expected, the increase in the number of visitors contributes to the increase of ESM. However, the effects of food business and education level still hold with the changes in visitor volumes.

We also tested the effect of newly established businesses using a different data source (Reference USA) and *Street Score* on the

**Table 1.** Change of ESM 2016–2018 (street-level).

| | Δ ESM | | | |
| --- | --- | --- | --- | --- |
| | (1) | (2) | (3) | (4) |
| Δ Pop Den | 0.004* | | 0.004* | 0.004* |
| | (0.002) | | (0.002) | (0.002) |
| Δ MH income | −0.001 | | −0.001 | −0.001 |
| | (0.002) | | (0.002) | (0.002) |
| Δ% bachelor | 0.012*** | | 0.012*** | 0.012*** |
| | (0.002) | | (0.002) | (0.002) |
| Δ Resi. diversity | 0.000 | | −0.000 | −0.000 |
| | (0.002) | | (0.002) | (0.002) |
| Δ food business | | 0.003** | 0.003** | 0.006*** |
| | | (0.001) | (0.001) | (0.002) |
| Δ food business× ESM 2016 | | | | −0.006** |
| | | | | (0.002) |
| Observations | 3865 | 3865 | 3865 | 3865 |
| $R^2$ | 0.5586 | 0.5559 | 0.5591 | 0.5598 |
| Fixed effect (county) | Yes | Yes | Yes | Yes |
| Spillover effect | Yes | Yes | Yes | Yes |
| Trend control | Yes | Yes | Yes | Yes |

OLS estimates on change of ESM from 2016 to 2018. Only streets with at least 20 unique visitors in both observation periods (84 days in each year) are included. Clustered standard errors in parentheses. # refers to count. Each street segment contains at least three POIs in year 2016. MH Income stands for Median Household Income. Resi. stands for residential. Extended table results in Supplemental Note and Table S3.
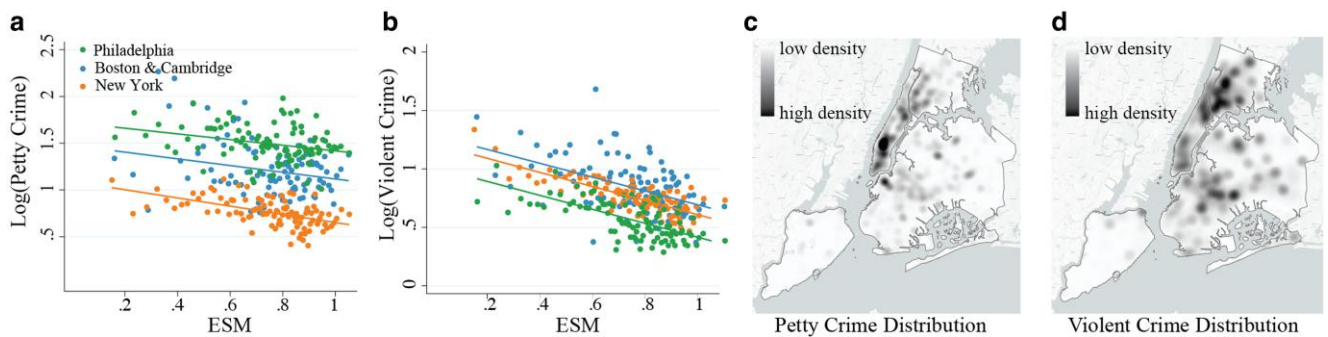***Coefficient significant at the 0.5% level, **at the 5% level, and *at the 10% level.



**Fig. 5.** ESM and crime. a) Binned scatter plot between ESM and log-transformed number of petty crimes (with the number of visitors by street segment controlled). b) Binned scatter plot between ESM and log-transformed number of violent crimes (with the number of visitors by street segment controlled). c) Kernel density illustration of petty crime density in New York City. d) Kernel density illustration of violent crime density in New York City (see Supplemental Fig. S9 for other cities).

changes of ESM. Supplementary Table S4 reports all results. The change of *Street Score* does not have a significant connection with the change of ESM. This is also potentially due to the fact that the change in urban appearance between the 2 years is relatively subtle. Consistent with the result of the change in the food business, the establishment of new business has a positive effect on the change of ESM.

## ESM and crime

One of the main concerns with dense cities is crime (19, 36). If a higher ESM indicates a higher chance for people with diverse backgrounds to meet each other, will it lead to more crime incidents? To understand the potential connection, we obtain crime reports from four sample cities within our study areas: New York City, Boston, Cambridge, and Philadelphia. Fig. 5a and b plots the relationships between crimes and the street-level ESM, conditioning on the number of visitors. We found that the number of petty and violent crimes (see Methods section for detailed definition) has a negative relationship with ESM. Conditioning on residential population, income level, number of visitors, residential diversity, and the number of POI, the street

segments with 1 SD higher ESM is associated with around 1% fewer violent crimes and pretty crime (see Supplemental Note 8 and Tables S7–S9 for the full results). This result implies that social mixing does not need to come at the price of more crimes. On the contrary, we can still create a socially mixed street environment with fewer crimes.

## Discussion

Many forces like gentrification, redlining, housing, or income inequality tend to segregate people in our cities. Curating a socially mixed urban environment is a common challenge for cities that aim at the overall goal of sustainability. Our study contributes to the current literature in this context from three perspectives. First, we investigate how the street sidewalk, as one of the most significant urban public spaces, can bring people from different income backgrounds together. The social mixing measured at street segment level rather than neighborhood or city represents the direct environment people will encounter throughout their day of life in cities. Second, we show that a city as a system has tools to leverage and curate a socially mixed environment in the urban activity space. Previous literature has largely focused on

**Table 2.** Summary statistics of the dataset per metropolitan area.

| Metropolitan area | #Devices | #Stays | #Street segment | #Census tract | #POIs | #Street images |
|---|---|---|---|---|---|---|
| Boston–Cambridge–Newton (2016) | 66k | 6.3M | 17k | 1007 | 14.4k | 0.23M |
| Boston–Cambridge–Newton (2018) | 174k | 12.2M | 17k | 1007 | – | 0.18M |
| New York–Jersey City | 210k | 42.2M | 77k | 4682 | 64.5k | 0.41M |
| Philadelphia–Camden–Wilmington | 112k | 14.8M | 26k | 1477 | 19k | 0.75M |

Statistics summary for three cities included. The number of devices only includes those that are identified with a home CBG. The number of stays only includes the stays that are not within a 50-m buffer of their associated home location.

using density as an instrument to foster diversity. Our cross-sectional models indicate that the residential social mixing can explain a large amount of ESM. This result implies that policy interventions such as mixed-income housing and affordable housing may have a profound impact on the level of diversity experienced by citizens beyond their residential areas. We also found that various types of venues might contribute to ESM at a different time of day—while coffee and tea shops contribute most to the ESM throughout the day, bars, restaurants, and retails contribute more ESM in the later afternoon. This result suggests that cities should consider both functional and temporal mix of venues when planning for a diverse urban environment. In addition, our longitudinal models reveal that by increasing the number of food businesses and attracting residents with higher education, the city could further improve street-level diversity. Lastly, we also found a potential connection between crime and ESM. Even though large cities with higher amount of visitors tend to have higher crime rates, we show that the streets with higher ESM tend to have lower crime counts when conditioning on the total number of visitors.

To this end, we also need to clarify a series of keywords in this study that might have been used differently in other literature. Our definition of ESM refers to the diverse income group of people one might encounter at a given street sidewalk segment. This definition stems from the overarching concept of "diversity" in today's urban planning practice, commonly describing land uses, architecture styles, street types, social classes, and skill sets. Similarly, the "density" in our study is also composed of different measures. Our study's core measure of density focuses on the number of visitors visiting a street segment. This measure is used as it is directly related to the ESM. In today's planning literature, scholars mostly use "density" to describe residential density. It is highly supported that diversity in cities necessitates a high density (18, 19). In this study, we use the residential density as a factor to test if it also contributes to the ESM and found that at our study scale (street segment), the streets with high residential density are less socially mixed. However, we should note that Jane Jacobs' observation of street "vitality" can still exist in residential street sidewalks with high residential density (2). Her observation of street vitality refers to a variety of activities happening on street sidewalks, such as playing soccer, conversing with neighbors, and eating and drinking.

Our study has several limitations. First, this study only discusses social mixing from the income aspect. Other forms of social mixing, measured from race, or occupation, might have a different presentation from income mixing. Our measure of social mixing uses proximity as a proxy, thus cannot determine if people staying at the same street segment are having meaningful interaction or not. In this regard, studies direct leveraging video footage in smaller urban spaces could further reveal the different kinds of interaction along the streets. Lastly, in this study, we used a longitudinal study to reveal the potential causal relationship

between the change of urban venues and ESM. However, as our data only have a two-year difference, factors such as residential income mixing and income level might not yet have changed. Future studies with longer time-span could further help identify the potential cause of ESM changes in cities, at a fine spatial scale.

## Methods
### Street segment

All street segments are downloaded from the OpenStreetMap through the python OSMnx package. Each street segment is defined as a segment of a street between two intersections. Each street segment contains an ID from the OpenStreetMap, a pair of $u$ and $v$ values representing the intersection node and the function type of the street segment. Duplicated geometries are removed from the original dataset. We also remove the major highway, primary links, secondary links, trunks, services streets, footpaths, steps, and slopes from the original dataset. (Note here that although we focus on the pedestrian network, footpaths are removed from the original dataset given its complexity and potential duplicates in a small space.) Only street segments with at least one POI within a 100-m buffer radius are included in the study. A total of 151,680 street segments from 3 cities are included in the study.

### Mobility data
*Attribution of stays to streets*

Each street segment is represented as a line in space. To attribute each stay to a street segment, we find the closest street segment for each stay. To avoid attributing a stay to a distant street segment, we choose only the street segment within a $d_{max} = 100$ m from each stay. If a stay is further than $d_{max}$ from any street segment, we discard it from the dataset. Fifty percent of the stays are within a 26.7-m radius from their closest street segment. The average distance of a stay to the closest street segment is around 31.2 m for all three cities. Distance is calculated based on each state's NAD83 state plane projection.[b]

*Street-level activities*

As we only focus on street-level activities, all stays that have a longer duration than 2 h are discarded from the dataset. Any stays within a $d_{home} = 50$ m from the identified home locations are also dropped from the study. The total number of stays and unique devices are shown in Table 2.

*Identifying home and economic status*

For each smartphone, we use its stays from 22:00 to 6:00 and spatially cluster them using the density-based spatial clustering of applications with noise (37) algorithm to detect the most likely cluster of stays each individual is located in during nighttime and early morning hours. We use 2 as the minimum number of

points per cluster and $\epsilon = 50$ m as the neighborhood. Then we join all detected cluster centers with each CBG geometry. We only consider individuals who were at the same CBG geometry for more than five nights in the observation period (3 months), and this CBG is considered as the home for this user. We use this CBG's median household income during the associated year to estimate the user's income level. This process leaves us to consider only 0.5 million users. Although mobile phone users are a large sample group, we admit that our data are still a sample of the true population. Robustness tests on population and income are included in Supplemental Note 2. Post-stratification was implemented to assure the representatives of the data in terms of income and population (see Supplemental Note 2).

## Measuring ESM
### Create income groups
We compare the median household income inferred from each individual's home CBG with the distribution of income in the metropolitan area so each CBG is assigned to a quartile of economic status within each metropolitan area. For each metropolitan area, the intervals of median household income for each economic group are different (see Supplementary Fig. S1).

### Street-level measure
To measure the ESM of each street $s$ in each city, we compute the proportion of total time spent at that street $s$ by each income quartile $q$ during the selected period $h$, $\tau_{qsh}$. Then we define $ESM_{hs}$ as the Shannon entropy (29) of each income group's activities at a given street segment $s$'s during a given time frame $h$:

$$ESM_{sh} = -\frac{1}{\log 4} \sum_{q=1}^{4} \tau_{qsh} \log(\tau_{qsh}), \qquad (1)$$

where $ESM_{hs}$ equals 0 when all users who visit the street $s$ in time period $h$ are from the same income group, while a larger value of the $ESM_{hs}$ means users from all four income groups spend a more equal amount of time visiting the street $s$ during period $h$. Only street segments with at least 20 users during a given period $h$ are included in the study to avoid severe small sample bias. The daily ESM only considers stays from 6 AM to 10 PM. The ESM at other periods is as specified in the paper.

### Census tract-level measure
Like street-level ESM, the CT-level ESM is the entropy of each income group's activities within a CT during a given time frame. Each stay is attributed to a CT through a spatial join process (see Supplementary Fig. S4 for the CT-level and street-level ESM comparison).

## POI data
### POI to street
POI data are from Foursquare (detail POI types are shown in Supplementary Table S6). We assign each POI to a street segment if it falls into a street segment's 100-m buffer. Each POI is also joint spatially with a CT that it falls into. The POI distribution within each city is shown in Supplementary Table S5.

### Change of business
The change in the food business is obtained from Boston's Hidden Restaurant. The data contain the restaurants, cafes, bars, and other food-related businesses that are closed or open in each month since 2007. For this study, as the mobility data cover

October to December in 2016 and October to December in 2018, we only select the restaurants that are either open or closed from 2017 January to 2018 September. The latitude and longitude of each restaurant were verified through google geocoding API. The chain stores are verified through Yelp. For the open and closed months for each recorded store, we also found similar results through the date of yelp reviews (see Supplementary Fig. S7 for change of business examples).

## Street score
To quantify the physical appearance of the built environment, we obtain 360° panorama GSV images of streetscapes through Google Maps API in all three study areas. Each panorama is associated with a unique identifier, latitude, longitude, month, and year of when the image was captured. We specify four angles to capture the full panorama of each street view location. To avoid the seasonal effect, we only keep images taken between April and October. GSVs taken from 2015 and 2016 are used in the cross-sectional study for the 2016 panel. GSVs taken from 2018 and 2019 were used for the 2018 panel. Moreover, images that were taken interior or highway only were excluded from the dataset. A total number of 1.5 million GSVs were used in the study (Table 2).

We measure the appearance of the built environment with a "Street Score," which indicates the perception of the safety of a GSV image. We use a deep learning model (31) pre-trained with a crowd-sourced dataset called Place Pulse, which contains millions of ratings on around 110,000 street view images from all over the world (24). The image diversity and rating consistency were evaluated by previous works (24, 38), indicating no significant bias depends on raters' cultural backgrounds in the dataset. We predicted the perception of safety for each image by ignoring the features of the sky, cars, and people in the dataset to minimize effects from time of day and other dynamic events (see Supplementary Note 6 and Fig. S8). The predicted continuous score ranges from 0 to 10, with 0 being the least safe-looking and 10 being the most safe-looking view. Then the "Street Score" for each CT and the street segment is the average score of all images associated with the CT and the street segment.

## Other data
Demographic data at the level of CBG and CT were obtained from the 5-year American Community Survey (ACS; 2012–2016 and 2014–2018).

### Residential income mixing
The residential income mixing is calculated using the same income group quartile per metropolitan area. To be consistent with the ESM calculation, at street level, we first buffer the street for 800 m and extract all CBGs that intersect with the street buffer. Then using the pre-assigned income quartile based on each CBG's median household income and population, we calculate the street-level residential mixing as the equation below:

$$R_c = -\frac{1}{\log 4} \sum_{q=1}^{4} n_{qc} \log(n_{qc}), \qquad (2)$$

where $n_{qc}$ is the population with the median household income level belonging to income quartile $q$. To test the robustness of this method, we also repeat the calculation by buffering from the street at 400 and 1000 m (see Supplementary Note 4.4).

### Crime incidents

The 2016 crime reports within the four sample cities are downloaded from each city's open data website. The original crime data come with crime primary types, crime incident date, and address. All four cities also provide crime incident locations' latitudes and longitudes except Cambridge, MA. We retrieve the latitude and longitude of crimes in Cambridge using the Google Map API geocoding service. Then we aggregated each crime incident to the street level by associated crimes to a street segment within a 30-m buffer distance. Two main types of crimes are separated from the original data. Violent crimes include rape, robbery, felony, or aggravated assault, and homicide or murder. Petty crimes include theft and larceny. Maps for all three cities are shown in Supplementary Fig. S9.

### Other

The population density, percentage of people with at least a bachelor's degree, and median household income are derived from the same CBGs to calculate the residential income diversity.

## Regression specification

### Cross-sectional model

We first specify the ordinary least-square model (OLS) to explain the ESM at the street level and CT level.

$$Y = \{Context\} + \{X\} + \epsilon, \tag{3}$$

$$Y = \{Context\} + \{X\} + \{Density\} + \epsilon, \tag{4}$$

where $Y$ is the estimated ESM of each experiment. $\{Context\}$ is a set of variables to control for the geographical context, including the segment length for the street-level experiment, the street segment's distance from the metropolitan center, and the overall sinuosity of the street segment shape (see Supplemental Note 4.1 for the method calculating sinuosity), land area size for the CT-level experiment, and county-level fixed effects. $\{X\}$ includes the variables we are interested to test: residential mixing, income, population density, venues count, and *Street Score* predicted from street view image data. The median household income comes from the ACS (5-year) survey corresponding to the mobility data's associated year. To account for the effect of density, we include a $\{Density\}$ term in Eq. 4, which stands for the total number of visitors. Finally, $\epsilon$ is the error term of the model.

### Spillover and network effects

Street segments are inevitably part of a larger network. The level of social mixing in one street segment could spill over through the network to its adjacent neighbors. To account for this potential spatial spillover effects, we estimate a variant of Eqs. 3 and 4 by

$$Y_i = \{Context\}_i + \{X\}_i + Y_{Spillovers,i} + \epsilon, \tag{5}$$

$$Y_i = \{Context\}_i + \{X\}_i + \{Density\}_i + Y_{Spillovers,i} + \epsilon, \tag{6}$$

where the $Y_{Spillovers,i}$ is the average ESM of segments that are connected with the street segment $i$. $Y_{Spillovers,i}$ is computed via $Y_{Spillovers,i} = \sum_i jWY_j$, where $W$ is the spatial weight matrix similar to Queen Continuity for polygon—we consider connected street segments as neighbors. Accordingly, we estimate Eqs. 5 and 6 via a G2SLS. The estimates are reported in Supplemental Table S1 (columns 5, 6, and 7). We found that the network spillover effect is an important determinant of the ESM level of a street segment. In particular, a 1 SD increase in the average neighborhood ESM is associated with a 0.89 SD increase in the segment

ESM. However, accounting for the spatial spillover effect does not change the qualitative results on residential diversity, the number of venues, and Street Scores. Robust standard errors are clustered at CBGs level.

### Difference-in-difference specification

To answer the question of which features contribute to the change of ESM, we specify the following equation:

$$\begin{aligned} \Delta Y = \{\Delta R\} + \{\Delta B\} + \{X\} + Y_{2016} \\ + Y_{2016} \times \{\Delta B\} \\ + \Delta Y_{spillover}, \end{aligned} \tag{7}$$

where $\Delta Y$ is the change of ESM from 2016 to 2018. $\{\Delta R\}$ is a set of demographic variables that change values from 2016 to 2018. The demographic data for 2018 are from ACS 2013–2018 survey. The $\{\Delta B\}$ is the change of food business aggregated at street level. $\{X\}$ includes a set of demographic variables in 2016 to control for the trend. To test the robustness, we also used the number of newly established businesses from the Reference USA 2017 data as the $\{\Delta B\}$ in additional tests. Similarly to Eqs. 5 and 6, this equation considers the network spillover effect of changes of ESM from 2016 to 2018 via the $\Delta Y_{spillover}$.

## Supplementary material

Supplementary material is available at *PNAS Nexus* online.

## Authors' contributions

Z.F. and E.M. designed the research; Z.F. performed the analysis; T.S., M.S., and F.Z. performed part of the analysis; Z.F. wrote the original manuscript; E.M, A.N., T.S., M.S., and F.Z. revised the manuscript; A.P. edited the manuscript. All authors reviewed the manuscript.

## Preprints

A preprint of this article is published at https://doi.org/10.48550/arXiv.2209.07041

## Data availability

The analysis was conducted using Python and Stata. Code to reproduce the main results in the figures from the aggregated data is publicly available on a GitHub repo https://github.com/brookefzy/social-mixing-street. For data exploratory purposes, the web map can be accessed via http://greatstreets.mit.edu/

# References

1 Jacobs AB. 1995. *Great Streets*. Cambridge: The MIT Press.
2 Jacobs J. 2020. The uses of sidewalks: safety. In: LeGates RT, Stout F, editors. I *The city reader*. 7th ed. London: Routledge. p. 114–118.
3 Glaeser E. 2012. *Triumph of the city: how our greatest invention makes us richer, smarter, greener, healthier, and happier*. New York: Penguin Group (USA).
4 Florida R. 2005. *Cities and the creative class*. New York: Routledge.
5 Florida R. 2006. The flight of the creative class: the new global competition for talent. *Liberal Educ*. 92(3):22–29.
6 Landry C. 2012. *The creative city: a toolkit for urban innovators*. 2nd ed. London: Earthscan.
7 Mayer SE. 2002. How economic segregation affects children's educational attainment. *Soc Forces*. 81(1):153–176.
8 Kramer MR, Hogue CR. 2009. Is segregation bad for your health? *Epidemiol Rev*. 31(1):178–194.
9 Athey S, Ferguson B, Gentzkow M, Schmidt T. 2021. Estimating experienced racial segregation in US cities using large-scale GPS data. *Proc Natl Acad Sci USA*. 118(46):e2026160118.
10 Moro E, Calacci D, Dong X, Pentland A. 2021. Mobility patterns are associated with experienced income segregation in large US cities. *Nat Commun*. 12(1):1–10.
11 Xu Y, Belyi A, Santi P, Ratti C. 2019. Quantifying segregation in an integrated urban physical-social space. *J R Soc Interface*. 16(160): 20190536.
12 Dong X, *et al*. 2020. Segregated interactions in urban and online space. *EPJ Data Sci*. 9(1):20.
13 Whyte WH. 1980. *The social life of small urban spaces*. Washington (DC): Conservation Foundation.
14 Gehl J. 2011. *Life between buildings: using public space*. Washington (DC): Island Press.
15 Cabrera JF, Najarian JC. 2013. Can new urbanism create diverse communities? *J Plan Educ Res*. 33(4):427–441.
16 Holden E, Norland IT. 2005. Three challenges for the compact city as a sustainable urban form: household consumption of energy and transport in eight residential areas in the greater Oslo region. *Urban Stud*. 42(12):2145–2166.
17 Wirth L. 1938. Urbanism as a way of life. *Am J Sociol*. 44(1):1–24.
18 Hill DR. 1988. Jane Jacobs' ideas on big, diverse cities: a review and commentary. *J Am Plan Assoc*. 54(3):302–314.
19 Moroni S. 2016. Urban density after Jane Jacobs: the crucial role of diversity and emergence. *City Territ Archit*. 3(1):1–8.
20 Calthorpe P, Van der Ryn S. 1986. *Sustainable communities: a new design synthesis for cities, suburbs and towns*. San Francisco, CA: Sierra Club.
21 Grant J. 2002. Mixed use in theory and practice: Canadian experience with implementing a planning principle. *J Am Plan Assoc*. 68(1):71–84.
22 Jacobs J. 1993. *The death and life of great American cities*. New York: Random House.
23 Branas CC, *et al*. 2018. Citywide cluster randomized trial to restore blighted vacant land and its effects on violence, crime, and fear. *Proc Natl Acad Sci USA*. 115(12):2946–2951.
24 Naik N, Philipoom J, Raskar R, Hidalgo C. 2014. Streetscore-predicting the perceived safety of one million streetscapes. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, Columbus, Ohio, p. 779–785.
25 Bogomolov A, *et al*. 2015. Moves on the street: classifying crime hotspots using aggregated anonymized data on people dynamics. *Big Data*. 3(3):148–158.
26 Bogomolov A, *et al*. 2014. Once upon a crime: towards crime prediction from demographics and mobile data. In: Proceedings of the 16th international conference on multimodal interaction, Istanbul, Turkey, p. 427–434.
27 Zhang F, Fan Z, Kang Yo, Hu Y, Ratti C. 2021. "perception bias": deciphering a mismatch between urban crime and perception of safety. *Landsc Urban Plan*. 207:104003.
28 Levy BL, Phillips NE, Sampson RJ. 2020. Triple disadvantage: neighborhood networks of everyday urban mobility and violence in US cities. *Am Sociol Rev*. 85(6):925–956.
29 Shannon CE. 1948. A mathematical theory of communication. *Bell Syst Tech J*. 27(3):379–423.
30 DeNadai M, *et al*. 2016. The death and life of great Italian cities: a mobile phone data perspective. In Proceedings of the 25th international conference on world wide web, Montreal, Canadap. 413–423.
31 Zhang F, *et al*. 2018. Measuring human perceptions of a large-scale urban region using machine learning. *Landsc Urban Plan*. 180:148–160.
32 Salesses MP. 2012. *Place Pulse: measuring the collaborative image of the city [PhD thesis]*. Cambridge: Massachusetts Institute of Technology.
33 Skogan W. 1986. Fear of crime and neighborhood change. *Crime Justice*. 8:203–229.
34 Kolko J. 2007. The determinants of gentrification. Available at SSRN 985714.
35 Lynch K. 1960. *The image of the city*. Vol. 11. Cambridge: MIT Press.
36 Glaeser EL, Sacerdote B. 1999. Why is there more crime in cities? *J Polit Econ*. 107(S6):S225–S258.
37 Ester M, Kriegel H-P, Sander J, Xu X. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In: kdd, Vol. 96. AAAI Press. p. 226–231.
38 Dubey A, Naik N, Parikh D, Raskar R, Hidalgo CA. 2016. *Deep learning the city: quantifying urban perception at a global scale*. In: 14th European conference on computer vision, Amsterdam, The Netherlands. Springer. p. 196–212.