








ARTICLE



<https://doi.org/10.1057/s41599-023-02570-5>

OPEN

Counterfactual mobility network embedding reveals prevalent accessibility gaps in U.S. cities

Yunke Zhang ^{1,7}, Fengli Xu^{1,7}, Lin Chen ², Yuan Yuan¹, James Evans ^{3,4}, Luis Bettencourt ^{4,5,6} & Yong Li ¹✉

Living in cities affords expanded access to various resources, infrastructures, and services at reduced travel costs, which improves social life and promotes systemic gains. However, recent research shows that urban dwellers also experience inequality in accessing urban facilities, which manifests in distinct travel and visitation patterns for residents with different demographic backgrounds. Here, we go beyond simple flawed correlation analysis and reveal prevalent accessibility gaps by quantifying the causal effects of resident demographics on mobility patterns extracted from U.S. residents' detailed interactions with millions of urban venues. Moreover, to efficiently reveal micro neighborhood-level accessibility gaps, we design a novel Counterfactual RANdom-walks-based Embedding (CRANE) method to learn continuous embedding vectors on urban mobility networks with confounding effects disentangled. Our analysis reveals significant income and racial gaps in mobility frequency and visitation rates to sports and education venues. Besides, bachelor's degree holders experience greater mobility reduction during the COVID-19 crisis. With extensive experiments on neighborhood-level accessibility prediction and visualizing accessibility gaps with embeddings vectors, we demonstrate that the counterfactual mobility network embeddings can improve the explanatory capacity and robustness of revealed accessibility gaps by extending them from aggregate statistics to individual neighborhoods and allowing for cross-city knowledge transfer. As such, urban mobility networks can reveal consistent accessibility gaps in the U.S., calling for urgent urban design policies to fill in the gaps.

¹Beijing National Research Center for Information Science and Technology (BNRist), Department of Electronic Engineering, Tsinghua University, Beijing, PR China. ²Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong SAR, PR China. ³Department of Sociology, University of Chicago, Chicago, IL, USA. ⁴Santa Fe Institute, Santa Fe, NM, USA. ⁵Mansueto Institute for Urban Innovation, University of Chicago, Chicago, IL, USA. ⁶Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA. ⁷These authors contributed equally: Yunke Zhang, Fengli Xu. ✉email: liyong07@tsinghua.edu.cn

Introduction

Urban dwellers live together in a compact environment but increasingly experience inequality in accessing urban venues and taking advantage of urban opportunities, which stems from the severe mobility gaps across gender, socioeconomic, and ethnic groups (Barbosa et al., 2021; Gauvin et al., 2020; Moro et al., 2021; Wang et al., 2018). Although urbanization improves venue accessibility and stimulates economic growth (Bettencourt, 2021), growing inequalities in urban mobility threaten the fabric that holds urban society together, affecting the livelihood and well-being of each resident (Glaeser et al., 2008). Inequality in urban mobility is entangled with multiple factors, including cultural experience and expectations (Wang et al., 2018), financial circumstance (Moro et al., 2021), and physical wellness (Althoff et al., 2017). As a result, measuring inequalities in urban mobility and dissecting its underlying factors represents a challenging task, but one that must be completed to meet UN sustainable development goals (SDGs) in reducing inequality and building resilient cities (Griggs et al., 2013).

There has been an increasing interest in uncovering mobility inequalities in urban spaces concerning various factors (Barbosa et al., 2021; Carra et al., 2016; Dueñas et al., 2021; Frias-Martinez et al., 2012; Gauvin et al., 2020; Graells-Garrido et al., 2020; Law, 1999; Lenormand et al., 2015; Macedo et al., 2022; Moro et al., 2021; Ng and Acker, 2018; Ryan et al., 2015; Sallis et al., 2018; Uteng, 2012). An important study (Law, 1999) defines the topic of “gender and daily mobility” and reviews studies on this from the early 70s. From a gendered perspective, a study in Chile finds that women not only visit fewer unique locations than men but also distribute their duration less equally (Gauvin et al., 2020). Mobility related to specific travel behaviors such as travel distance and the number of trips has been examined by gender in eight cities in multiple countries (Ng and Acker, 2018), with findings that women tend to travel shorter distances than men. Much research has investigated the relationship between other demographic or socioeconomic features and urban mobility. Both higher socioeconomic status (Frias-Martinez et al., 2012) and younger age (Lenormand et al. 2015) strongly correlate with a larger mobility range. A study in South America finds that middle-class travelers exhibit the most diverse mobility patterns, while the lower classes manifest limited spatial exploration (Macedo et al., 2022). Regarding the response to the COVID-19 pandemic, poorer populations show lower reductions in mobility level (Dueñas et al., 2021). Moreover, planned spatial factors play an important role in urban mobility, with previous studies revealing that improved access to public transportation infrastructure (Ryan et al., 2015) and well-designed neighborhood walkability (Sallis et al., 2018) can enhance mobility within urban areas.

Existing studies providing evidence on inequality in urban mobility, however, remain woefully inadequate due to limitations in datasets. Many efforts rely on small-scale survey data that is expensive to collect and difficult to scale up (Sharkey and Elwert, 2011). Some research attempts large-scale estimation via hypothetical models defined on spatial proximity (Hasthanasombat and Mascolo, 2019; Saxon, 2021), but this inevitably leads to bias in analytic results. Recent burgeoning precise mobility data provide the chance to compensate for the above deficiencies (Kadar et al., 2020; Moro et al., 2021). Here we measure accessibility gaps based on large-scale mobility data provided by SafeGraph company. SafeGraph curates mobile device users’ traces to obtain fine-grained visitations to urban facilities represented as Points-of-Interest (POIs). Covering over 35 million residents’ visits to more than 4 million POI in the U.S. from 2019 to 2020, the large-scale and accurate mobility dataset can sufficiently measure gaps in various mobility patterns that depict a resident’s capability to

exploit urban opportunities. Beyond data limitations, previous works also suffer from research methodologies. Most of them rely on correlation analysis between mobility patterns and a few specific demographic factors (Barbosa et al., 2021; Gauvin et al., 2020; Wang et al., 2018), which cannot control for confounding effects in urban mobility and may lead to inaccurate conclusions. Findings from correlation analysis are often based on aggregate statistics and provide limited insights into the micro-level causes of accessibility gaps. Therefore, it is important to tease apart potential confounding factors and probe micro-level gaps.

In this study, We present a novel framework to quantify accessibility gaps in six metropolitan areas of the US. We employ a propensity score matching (PSM) method that estimates the causal effects of demographic features on mobility patterns. Specifically, we look at the causal effect of gender, race, income, physical disability, and education background on *mobility frequency* and *urban facility accessibility* across different neighborhoods, which quantify the number of movements per person and the likelihood of accessing different urban facilities, respectively. Furthermore, we design a novel Counterfactual RANdom-walks-based Embedding (CRANE) to learn representations for micro-level accessibility gaps on urban mobility networks. Specifically, we use random-walk sampling to efficiently assess the empirically *observed* association between demographic and urban facility accessibility and the *alternative* outcomes in counterfactual scenarios where the demographic does not have a causal impact. Drawing on the difference between the *observed* and *alternative* outcomes, our CRANE method can efficiently approximate the causal inference result by treating them as positive and negative samples in representation learning, respectively. The representations for neighborhoods and POIs allow us to go beyond aggregate statistics and extend our analysis to their micro-level behaviors, enriching our understanding of the causes underlying accessibility inequality.

We observe interesting findings regarding mobility inequality in the US. Neighborhoods with higher average income and a higher portion of the white population consistently present higher mobility frequency across different cities, which suggests these sub-populations have superior access to urban facilities. Neighborhoods with a higher portion of bachelor’s degree holders consistently manifest greater mobility reductions in response to the COVID-19 pandemic, which implies that these neighborhoods can afford fewer outdoor activities, likely from elasticity regarding their work format. Sports and education venues are more likely visited by white and higher-income populations. Extensive experiments demonstrate that neighborhood and POI embeddings learned from counterfactual random walks on urban mobility networks can improve the performance of accessibility prediction compared to traditional correlation-based methods. A case study on the POI level illustrates that our method successfully disentangles confounding effects between neighborhood demographics.

Methods

Demographic dataset. We collect demographic data from the 2019 U.S. Census Bureau’s American Community Survey 5-year Estimates (ACS) as the independent variable of mobility patterns. In this dataset, features are reported at the level of census block groups (CBGs), the smallest geographical unit for publicly available census data. In this study, we use CBGs as a proxy for neighborhoods. We focus on the six most populated metropolitan statistical areas (MSAs) in the U.S.: New York-Newark-Jersey City, NY-NJ-PA (hereby referred to as the “New York” MSA, “NY” for short), Los Angeles-Long Beach-Anaheim, CA (“Los Angeles”,

Table 1 Basic statistics of ACS and SafeGraph datasets in 6 most populated US MSAs.

MSA	N_{CBG}	\bar{N}_p	\bar{N}_H	Female%	White%	Bachelor%	Income	Disability%	\bar{N}_{M2019}	\bar{N}_{M2020}	Δ_M	Art%	Sports%	Edu%	Health%
NY	14,308	1352.56	482.54	51.71%	58.09%	37.40%	95726	21.43%	778.83	453.16	325.66	2.41	6.37	6.53	4.41
LA	8239	1600.50	521.57	50.71%	55.02%	31.01%	87253	21.51%	740.58	360.28	380.32	1.91	6.01	6.51	3.68
Chi	6584	1446.81	526.05	51.09%	66.31%	35.15%	85475	20.97%	759.44	535.09	224.35	1.45	6.50	8.60	4.38
Dal	4128	1670.11	587.91	50.83%	69.49%	31.97%	83416	20.64%	1145.52	877.96	267.56	1.07	4.84	8.90	3.99
Hou	3017	2148.69	737.10	50.32%	65.69%	30.02%	85547	20.96%	1029.18	762.67	266.51	0.88	4.95	8.15	3.62
DC	3575	1685.28	604.48	51.18%	55.17%	48.50%	118040	17.71%	676.75	383.81	292.95	1.33	5.84	7.01	3.90

“LA”), Chicago-Naperville-Elgin, IL-IN-WI (“Chicago”, “Chi”), Dallas-Fort Worth-Arlington, TX (“Dallas”, “Dal”), Houston-The Woodlands-Sugar Land, TX (“Houston”, “Hou”), Washington-Arlington-Alexandria, DC-VA-MD-WV (“Washington DC”, “DC”). The numbers of neighborhoods in each MSA are listed as N_{CBG} in Table 1.

Table 1 gives an overview of the average population \bar{N}_p and the number of households \bar{N}_H of neighborhoods in each MSA. We extract the following demographic features as potential contributors to urban accessibility gaps: (1) **Female ratio** is the proportion of female residents in each neighborhood, which is used to analyze the potential gender gap in urban mobility (Gauvin et al., 2020). (2) **White ratio** is the neighborhood’s proportion of white population, which corresponds to the previous finding that the racial minorities experience social isolation when traveling in the city (Wang et al., 2018). (3) **Bachelor ratio** is the proportion of people over 25 years that have a bachelor’s degree or higher, reflecting the real-world observation that people with different educational backgrounds are likely to exhibit different mobility patterns. (4) **Average income** is the household average income in the past 12 months, corresponding to diverse mobility patterns associated with economic status (Macedo et al., 2022; Šćepanović et al., 2015). (5) **Disability ratio** is the proportion of households with at least one disabled resident whose mobility capability is constrained by the quality of urban infrastructure (Saha et al., 2021). The average values of the above neighborhood feature in each MSA are reported in Table 1.

Mobility dataset. To reflect the heterogeneity in mobility patterns across the city, we utilize SafeGraph’s Patterns dataset (<https://docs.safegraph.com/docs/monthly-patterns>) and Core Places dataset (<https://docs.safegraph.com/docs/core-places>). By tracking GPS-equipped mobile devices under consent, the Patterns dataset aggregates visit counts from CBGs to POIs each month, providing fine-grained mobility records of urban dwellers and detailed information about POIs. It is worth mentioning that workers at POIs are excluded from the visit counts, reflecting residents’ subjective capability of accessing urban venues instead of objective requirements. Across the six MSAs, the dataset records 155 million visits paid to 758 thousand POIs every month on average. The Core Places dataset identifies each POI’s category under the North American Industry Classification Systems (NAICS).

Combining the Patterns and Core Places datasets, we can quantify mobility patterns associated with each neighborhood. In this study, we adopt Amartya Sen’s capability analysis framework (Sen, 1980), where equality is defined as a person’s basic capability of being able to do certain basic things in the environment, e.g. move or be clothed. From the perspective of mobility, the capability of accessing and visiting more urban venues is often linked with higher social status (Barbosa et al., 2021; Chang et al., 2021; Chen et al., 2022; Lenormand et al., 2015; Wang et al., 2011; Xu et al., 2018). Thus, we are devoted to revealing the gap in two measures of urban mobility capability, namely **mobility frequency** (N_M) and **urban facility accessibility**.

First, to depict aggregate mobility capability for each neighborhood, we calculate its **mobility frequency** (N_M) as follows. We sum the neighborhood’s visits over POIs located in the MSA in three months (from July to September) to obtain the total number of visits made by its residents and then normalize it by the neighborhood’s population to obtain the average number of visits per 100 residents. To study the mobility patterns influenced by the COVID-19 pandemic, we further calculate the mobility reduction Δ_M by comparing N_M in 2019 and 2020. The average mobility frequencies \bar{N}_M and mobility reductions Δ_M for

Table 2 The correlation between demographic feature and mobility behavior in New York metropolitan statistical area.

	Female%	White%	Bachelor%	Income	Disability%
N_{M2019}	-0.037***	0.111***	0.054***	0.132***	-0.003
Δ_M	-0.025**	-0.067***	0.076***	0.035***	-0.040***
Art%	0.095***	-0.265***	-0.052***	-0.205***	0.073***
Sports %	-0.040***	0.520***	0.673***	0.643***	-0.289***
Education %	-0.026**	0.336***	0.117***	0.370***	-0.017*
Health %	0.061***	0.020*	-0.140***	-0.076***	0.133***

Significance level: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Bold text indicates that the corresponding correlation coefficient has the opposite sign to the treatment effect.

neighborhoods in each MSA are listed in Table 1. We can observe a consistent reduction in all analyzed MSAs, which reflects the overall impact of social distancing and stay-at-home orders during the pandemic.

Second, we calculate each neighborhood's **urban facility accessibility**, which is the proportion of its visits to a specific category of POIs to its total visits N_M . This pattern reflects urban dwellers' access to urban facilities. Specifically, we focus on four representative categories, i.e., Art and Recreation, Sports, Education, and Health. The average accessibility to four categories in each MSA in 2019 is listed in Table 1 as "Art%", "Sports%", "Edu%", and "Health%".

Preliminary analysis. We calculate the mobility patterns of each neighborhood in the six most populated US metropolitan statistical areas by the SafeGraph monthly patterns datasets covering July through September in the years of 2019 and 2020. A simple approach to measuring mobility inequality in an urban area is to calculate the correlation between a neighborhood's demographic features and its mobility patterns (See Supplementary Fig. S1 for the spatial distribution of mobility patterns and demographic features in New York). Table 2 shows correlations between mobility patterns and demographic features and their corresponding significance levels (two-sided p -values) in New York MSA as an example (See correlation tables of other MSAs in Supplementary Table S1–5). Significant correlations indicate mobility patterns vary substantially with the neighborhood's demographic features, indicating potential mobility inequalities. For example, neighborhoods with a higher bachelor ratio are correlated with higher mobility frequency. However, as observed in Supplementary Fig. S1, the *bachelor ratio* is also positively correlated with *average income*, which raises the concern of confounding: the correlation between *bachelor ratio* and mobility frequency might be sustained by their common correlations with *average income*. To address this concern, our subsequent analysis employs PSM to mitigate confounding bias and assess the impact of each feature on mobility patterns independently.

Propensity score matching. We adopt the PSM method, a *matching* method widely used in the causal inference literature (Rosenbaum and Rubin, 1983), to estimate the causal effect of demographic features on heterogeneous urban mobility patterns. The procedure of this method is illustrated in Fig. 1.

The core concept is to account for potential confounding factors denoted as *covariates* \mathbf{X} by matching each neighborhood with neighborhoods that share similar \mathbf{X} values but differ in a specific relevant demographic feature, referred to as the *treatment* T (Fig. 1A). The PSM procedure disentangles the influence of confounding effects by creating a *treatment* group comprising neighborhoods with comparable *covariates* \mathbf{X} and attributes the variations in mobility pattern within this group to differences in the *treatment* level, i.e., differences in the relevant demographic feature. However, exact matching based on high

dimensional *covariates* can be computationally expensive and often leads to sparsity issues. To address this, we employ a propensity score function $b(\mathbf{X})$ to map *covariates* into a scalar propensity score. Specifically, we first discretize the *treatment* T into five equally-sized bins, defined as *treatment* levels $L(T)$. Subsequently, we fit an ordinal regression model (McCullagh, 1980) to estimate propensity scores,

$$\log \frac{P(L(T) \leq L^*)}{P(L(T) > L^*)} = \theta_{L^*} - \mathbf{w}^T \mathbf{X}, \quad (1)$$

where L^* represents the *treatment* level, and θ_{L^*} along with \mathbf{w} are learnable model parameters. The ordinal regression model satisfies the condition that *treatment* level $L(T)$ and *covariates* \mathbf{X} are independent given $b(\mathbf{X}) = \mathbf{w}^T \mathbf{X}$. Therefore, We calculate $\hat{\mathbf{w}}^T \mathbf{X}$ as the estimated propensity score for each neighborhood, with $\hat{\mathbf{w}}$ being the model's fitted parameter. The distribution of \mathbf{X} conditioned on $b(\mathbf{X})$ tends to be similar across different values of $L(T)$. As a result, the estimated propensity scores can be employed as matching criteria. To perform matching, we define the distance between two neighborhoods as the disparity in estimated propensity scores divided by the difference in $L(T)$. Each neighborhood is then matched with its nearest neighbor using this distance metric.

Once we have established pairs of matched neighborhoods denoted as $M = \{(i, j)\}$, we estimate the average treatment effect of T on a specific mobility behavior Y . This estimation is based on the expected change in Y (which can represent mobility frequency or urban facility accessibility) when the treatment level $L(T)$ is increased by one unit of level,

$$\text{Average Treatment Effect} = \frac{1}{|M|} \sum_{(i,j) \in M} \frac{Y_i - Y_j}{L(T)_i - L(T)_j}. \quad (2)$$

A critical prerequisite for estimating unbiased treatment effects of demographic features on mobility patterns is the selection of *covariates*. In the PSM procedure, for each demographic feature as the treatment, we choose other demographic features along with the ratio of residents younger than 20 years and older than 60 years as *covariates* for the following reasons. First, all demographic features can potentially influence urban mobility. Previous studies have revealed disparities in urban mobility based on gender (Gauvin et al., 2020), race (Wang et al., 2018), economic status (Macedo et al., 2022), and disability (Saha et al., 2021). Additionally, differences in mobility patterns among different age groups are commonly observed in urban environments. Therefore, we consider all demographic features and age distribution as potential *covariates*. Second, an important characteristic of the urban environment is circular causality (Bettencourt, 2021). Demographic features at the CBG level also exhibit this characteristic. For example, CBGs with a higher percentage of bachelor's degree holders may attract residents with higher-paying careers, resulting in a higher average income for that CBG. Conversely, CBGs with better economic conditions may be more appealing to highly educated residents, leading to an

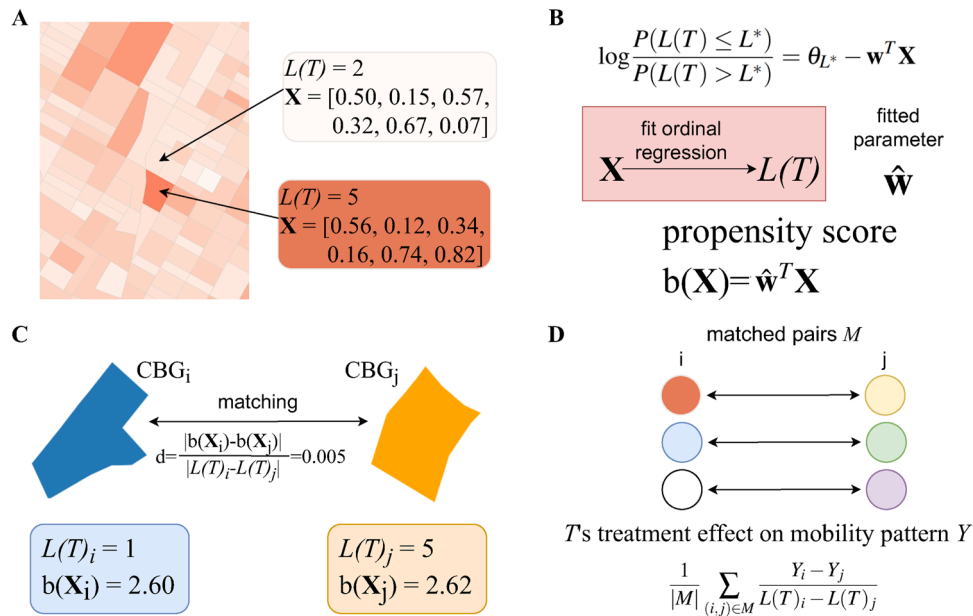


Fig. 1 A schematic representation of the propensity score matching method. A Treatment and covariates of urban CBGs. Treatment is stratified into 5 levels, denoted as $L(T)$. **B** Estimation of propensity scores by fitting an ordinal regression on $L(T)$ by covariates. **C** Matching two closest CBGs according to their distance determined as the difference in propensity scores divided by the difference in $L(T)$. **D** Calculating average treatment effect from matched CBG pairs M .

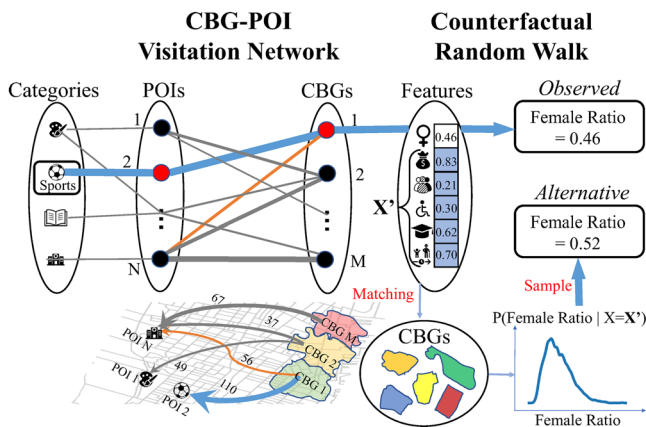


Fig. 2 Illustration of counterfactual random walks on urban mobility network. Starting from a POI category, a POI is sampled based on its total visitation frequency and an observed neighborhood is sampled based on its normalized visitation frequency to the sampled POI. An alternative neighborhood is sampled from all neighborhoods that have covariates identical to the observed neighborhood.

increased bachelor ratio. Therefore, for each treatment variable T , we include the other demographic features along with the age distribution within the CBG as covariates represented by vector \mathbf{X} . In each MSA, we employ PSM to estimate the average treatment effects of five demographic features on mobility patterns, including mobility frequency in 2019, its reduction in 2020, and urban facility accessibility to four POI categories.

Counterfactual random walk on urban mobility network. While PSM estimation is effective, it has limitations as it can only produce aggregate statistics at the MSA level, i.e., average treatment effects. These statistics provide limited insights into the micro-level mechanisms of mobility inequality and are not easily applicable to downstream applications or analyses. Drawing

inspiration from recent studies that uncover gender disparities using word embeddings (Garg et al., 2018), we develop a novel representation learning algorithm to capture micro-level disparities in facility accessibility using continuous embedding vectors. Our approach involves conducting counterfactual random walks on an urban mobility network to assess the empirical association between a specific demographic feature and the accessibility of a particular category of POI. We compare these associations with alternative outcomes that would occur if the demographic feature did not causally affect access to the POI category. This process allows us to update corresponding embeddings that preserve real-world gaps in accessibility data.

Specifically, we first construct an urban mobility network that consists of four types of heterogeneous nodes, representing POI categories, POIs, neighborhoods (CBGs), and demographic features, shown in Fig. 2. There are three types of edges in the constructed network, i.e., POI-category edges and neighborhood-demographic edges that connect each POI and neighborhood to its category and demographic feature, respectively, and POI-neighborhood edges weighted by visitation frequency in mobility data. In each random walk, we first select a POI category Q and sample a specific POI P_i based on its total visitation frequency. Then, we sample a neighborhood C_o with probability proportional to the edge weights w_{io} between P_i and C_o , normalized by the out-degree of C_o , i.e., $Prob(C_o|P_i) \sim w_{io}/\sum_k w_{ko}$. It captures the likelihood for neighborhood C_o to visit that POI, i.e., the accessibility of that urban facilities. Finally, we look up the value of the interested demographic feature of C_o as the *observed outcome*. Following this procedure, we can sample a $Q \rightarrow P_i \rightarrow C_o \rightarrow o$ path that iteratively connects the POI category, POI, neighborhood, and its demographic feature. It serves as a sample of the empirically observed accessibility from the interested demographic feature to a POI category.

Aside from the *observed outcome*, we sample a neighborhood C_a as the *alternative outcome* in the counterfactual scenario that the interested demographic feature does not have a causal effect on the access to that POI category. Specifically, we mimic the

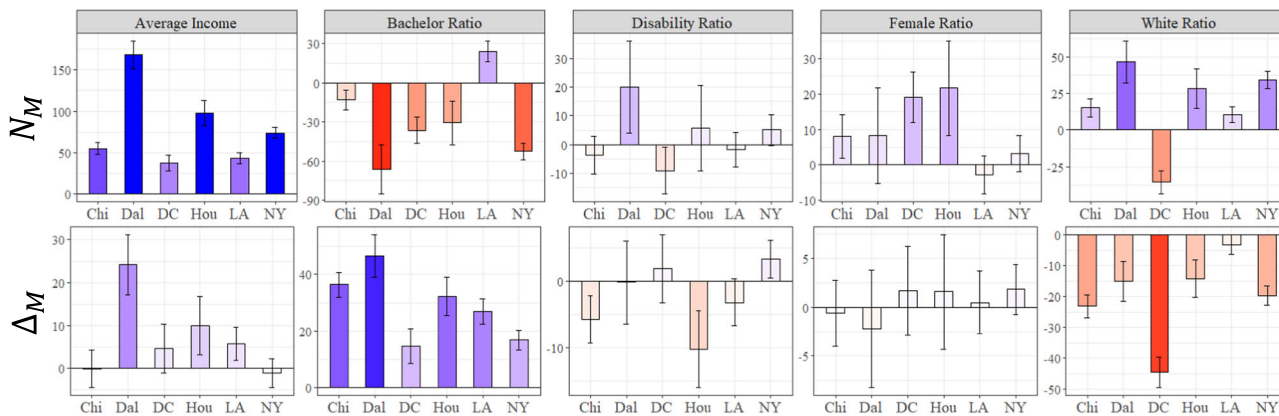


Fig. 3 The treatment effects on mobility frequency in 2019 and mobility reduction during COVID-19 pandemic. Each panel denotes the treatment effects of one demographic feature on a mobility pattern in six MSAs. Blue and red bars indicate positive and negative effects, respectively. Whiskers correspond to the 95% confidence intervals.

matching process in PSM to sample a neighborhood C_a with the same stratified *covariates* as C_o and take its demographic feature as the *alternative outcome*. For instance, as shown in Fig. 2, we randomly sample a neighborhood from the pool that has the same stratified *covariates* as C_o and take its female ratio as the *alternative outcome*. Intuitively, the *alternative outcome* assesses the counterfactual scenario that the demographic feature has no causal impact on the accessibility to specific urban facilities, and hence its distribution is solely conditioned on the *covariates*. Therefore, if the *observed outcome* and *alternative outcome* follow significantly different distributions, it indicates the counterfactual scenario is unlikely and the demographic feature probably has a causal effect, which provides a signal to update the node embeddings.

Network embedding for mobility inequality. On top of the sampled random walks, we seek to preserve the difference between *observed* and *alternative outcomes* in the embedding space to capture the gaps in urban facility accessibility. Specifically, the *alternative outcomes* serve as the negative samples that the examined demographic feature does not have a causal impact. Therefore, in each random walk, we maximize the embedding similarity for the co-occurrences of *observed outcome*, observed neighborhood, POI, and POI category, and minimize the embedding similarity for the co-occurrences of *alternative outcome*, alternative neighborhood, POI and POI category.

In each MSA, we learn an embedding vector for each of the following entities: four POI categories, all POIs, all neighborhoods, and all treatment levels of five neighborhood demographic features. *Observed outcomes* and *alternative outcomes* are represented as the vector of its corresponding treatment level $L(T)$. We train embedding vectors to minimize the following loss function:

$$\begin{aligned}
 Loss = & \sum_{\{(c, P, C_o, C_a, o, a)\}} -\log \sigma(\mathbf{v}_c^T \cdot \mathbf{v}_o) - \log \sigma(-\mathbf{v}_c^T \cdot \mathbf{v}_a) \\
 & - \log \sigma(\mathbf{v}_c^T \cdot \mathbf{v}_{C_o}) - \log \sigma(-\mathbf{v}_c^T \cdot \mathbf{v}_{C_a}) - \log \sigma(\mathbf{v}_P^T \cdot \mathbf{v}_o) \\
 & - \log \sigma(-\mathbf{v}_P^T \cdot \mathbf{v}_a) - \log \sigma(\mathbf{v}_P^T \cdot \mathbf{v}_{C_o}) - \log \sigma(-\mathbf{v}_P^T \cdot \mathbf{v}_{C_a}),
 \end{aligned}
 \tag{3}$$

where $\{(c, P, C_o, C_a, o, a)\}$ is the set containing the sampled category, POI, neighborhood, alternative neighborhood, *observed outcome*, and *alternative outcome* walks, $\mathbf{v}_c, \mathbf{v}_P, \mathbf{v}_{C_o}$ are the POI category, POI, and neighborhood embedding vectors, $\sigma(\cdot)$ is the sigmoid function, \mathbf{v}_o and \mathbf{v}_a are embedding vectors of *observed outcome* and *alternative outcome*'s treatment levels, respectively.

Regularization terms are added to the loss function to achieve spatial smoothness of neighborhood embeddings and continuity in demographic feature embeddings. According to the First Law of Geography (Miller, 2004), nearby things are more related to one another. Two nearby neighborhoods should also be adjacent to our embedding space. The spatial regularization term is the weighted sum of squared distances between adjacent neighborhoods in the embedding space. Weights are defined as $\exp(-\frac{d^2}{2\sigma_s^2})/(2\pi\sigma_s^2)$ for each pair of neighborhoods with distance d less than a threshold σ_s . The demographic regularization term is the sum of squared distances between adjacent $L(T)$'s in the embedding space, e.g. the lowest and second-lowest level of *average income*. We use the Adam optimizer (Kingma and Ba, 2015) to learn 64-dimensional embedding vectors for each category, POI, neighborhood, and $L(T)$ to minimize the loss function (3). The spatial threshold σ_s is set to 2.5 kilometers, and the strength of regularization is 0.0001 and 0.01 for spatial and demographic feature continuity. See Supplementary Note S1.2 for implementation details and complexity analysis of the algorithm.

Results

Consistent gaps in urban accessibility

Mobility frequency. The first row in Fig. 3 shows the estimated treatment effects of demographic features on mobility frequency in each MSA. Each panel represents the effects of one demographic feature on N_M in six MSAs. Positive and negative effects are indicated by blue and red columns, respectively. Whiskers represent the 95% confidence interval of the treatment effect. Significant gaps are indicated by dark columns with a confidence interval on one side of the x-axis. According to Amartya Sen's capability analysis framework, increased mobility frequency signifies an enhanced capability to access and visit urban facilities. Negative treatment effects on mobility frequency imply reduced capability as the demographic feature increases within the neighborhood.

From our analysis of treatment effects, we summarize three key observations of mobility frequency gaps. First, neighborhoods with higher average income have higher mobility frequency. In all MSAs, treatment effects are over 30, implying that a neighborhood with one unit higher $L(income)$ will have approximately 30 more visits per 100 residents in three months. This is consistent with previous findings that low-income residents have lower mobility rates (Pucher and Renne, 2003), revealing an unignorable cost of moving around the city. Second, in terms of education, neighborhoods with a higher proportion of bachelor's

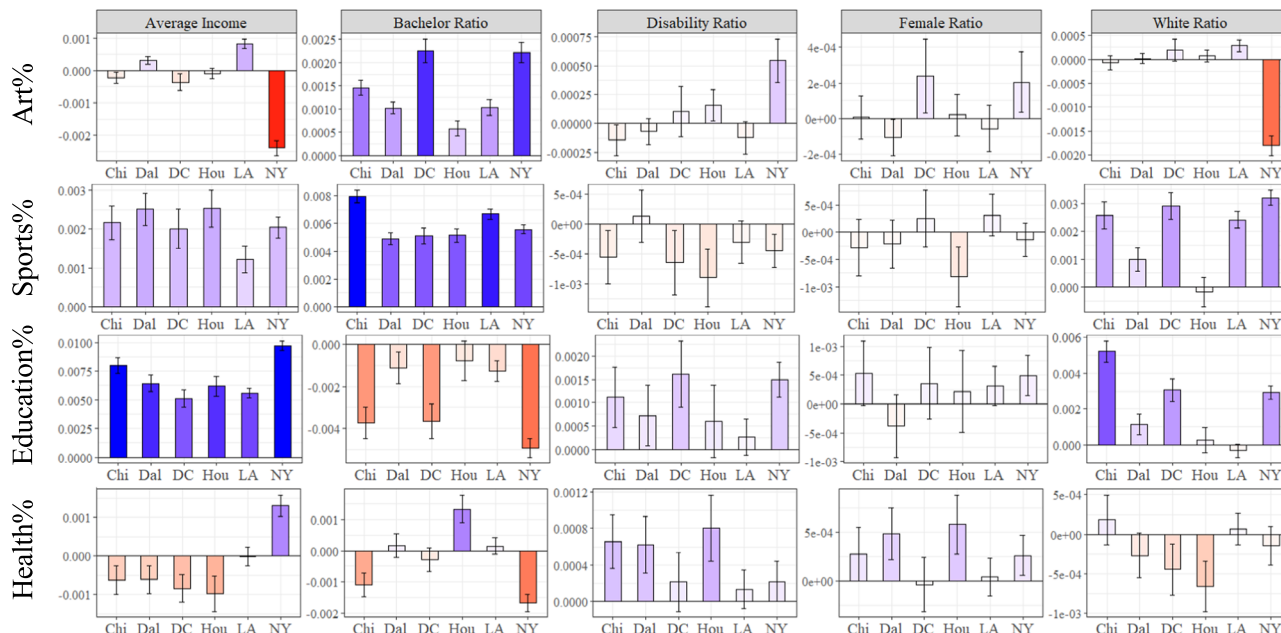


Fig. 4 The treatment effects on the accessibility to four POI categories. Each panel denotes the treatment effects of one demographic feature on the accessibility to a POI category in six MSAs. Blue and red bars indicate positive and negative effects, respectively. Whiskers correspond to the 95% confidence intervals.

degree holders demonstrate significantly lower mobility frequencies in most MSAs. This negative effect is opposite to the positive correlation reported in Table 2, underscoring the necessity of controlling for confounding demographic features. This difference implies that among individuals with similar income levels and other demographic characteristics, highly educated individuals access urban facilities less frequently. Third, neighborhoods with a higher white ratio have significantly higher mobility frequencies in most MSAs, which means that the white population is clearly advantaged in accessing urban facilities. This observation reveals the superior access to urban essential facilities of the white population in these MSAs.

Mobility reduction during the COVID-19 pandemic. Since the COVID-19 outbreak, authorities have issued mobility control measures such as social distancing and stay-at-home orders to request that urban dwellers reduce unnecessary travel. Although we observe a consistent reduction in the overall mobility frequency, such reduction is not experienced equally by all. To measure this gap, we show the treatment effect of neighborhood demographics on the reduction of mobility frequency in 2020 in the second row of Fig. 3, where blue columns indicate that a higher $L(T)$ of the demographic feature results in a larger mobility reduction.

We summarize two key observations as follows. First, we find that neighborhoods with a higher bachelor ratio have greater mobility reduction in all six MSAs. One explanation is that highly-educated residents can more easily transition to telework due to the nature of their occupations, e.g., working on personal computers (Lund et al., 2020). In contrast, people with inferior educational backgrounds are more likely to be essential workers whose work cannot be performed remotely, such as building cleaning and equipment maintenance. As a result, when public health crises force people to reduce mobility, inequalities faced by people of different educational backgrounds become exacerbated.

Second, we find that residents with different ethnicities are unequally affected by the pandemic. Specifically, neighborhoods with a higher white ratio have less mobility reduction in all 6

MSAs, which widens the pre-existing accessibility gap among ethnic groups. This is likely because white populations have superior access to urban facilities and opportunities, making them more resilient to pandemic shocks.

Urban facility accessibility. The treatment effects of neighborhood demographics on residents’ accessibility to four venue categories - Art & Recreation, Sports, Education, and Health in the six most populated MSAs are shown in Fig. 4.

Art and recreational venues are more accessed by neighborhoods with a higher bachelor ratio. Treatment effects are over 0.2% in New York and Washington DC. This reveals a difference in utilizing urban facilities of art and culture, with highly educated residents more likely to visit these venues. Sports facilities are more accessed by neighborhoods with higher average income, bachelor ratio, white ratio, and lower disability rates. This suggests that disabled residents have very limited access to sports facilities.

Meanwhile, uneven access is also prominently associated with income, race, and educational background. Educational POIs are more accessed by neighborhoods with higher average income, disability ratio, white ratio, and lower bachelor ratio. This indicates that wealthy and white populations are most advantaged in accessing educational facilities. Health services are more accessed by neighborhoods with higher disability rates and female ratios. This is reasonable as people with disabilities seek health services and rehabilitation more frequently and females carry more healthcare burdens. Surprisingly, neighborhoods with lower average income also pay a higher ratio of visits to health services in Chicago, Dallas, Houston, and Washington DC. It is likely because wealthier people resort more often to doctor home visits for care. This also reflects discrepancies in health status and healthcare needs, as wealthier people are also more likely to maintain better health with reduced onsite healthcare needs.

Besides identifying notable inequalities in mobility frequency and urban facility accessibility, we find interesting results showing how the PSM procedure alleviates confounding effects by balancing covariates. For instance, *bachelor ratio* is negatively

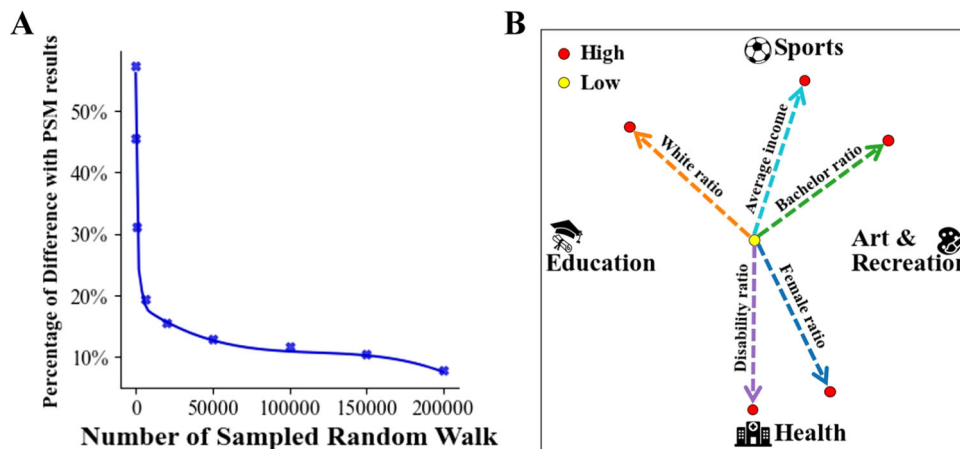


Fig. 5 Demonstration of the convergence and effectiveness of CRANE. **A** Convergence with sample size. The percentage of difference with PSM results decreases with the number of sampled random walks on the urban mobility network and is reduced under 10% when the sample size is over 200,000. **B** 2D visualization of learned embeddings. Four category embeddings are fixed to be projected onto (0,1), (1,0), (0,-1), and (-1,0). Dashed arrows represent the directions pointing to the highest $L(T)$ from the lowest $L(T)$'s embedding in the projected 2D space of each demographic feature.

Table 3 Regression coefficients between each demographic feature and its proximity with POI categories in the embedding space.

MSAs		New York				Los Angeles				Chicago			
Demographic	Category	Art%	Sports%	Edu%	Health%	Art%	Sports%	Edu%	Health%	Art%	Sports%	Edu%	Health%
Female%		0.023	0.018	0.030	0.003	0.014	0.005	0.018	-0.001	0.000	-0.002	-0.001	0.014
White%		-0.050	0.036	0.021	0.007	0.015	0.024	-0.001	-0.006	-0.005	0.041	0.036	-0.006
Bachelor%		0.033	0.031	-0.027	-0.017	0.044	0.047	-0.004	0.000	0.039	0.054	-0.010	-0.016
Income		-0.031	0.018	0.032	0.001	0.025	0.002	0.025	-0.011	0.011	0.016	0.016	-0.014
Disability%		0.025	0.002	0.016	0.016	-0.024	-0.016	0.001	-0.000	-0.026	-0.013	0.002	0.001
MSAs		Dallas				Houston				Washington DC			
Demographic	Category	Art%	Sports%	Edu%	Health%	Art%	Sports%	Edu%	Health%	Art%	Sports%	Edu%	Health%
Female%		-0.014	0.001	-0.010	0.011	-0.002	0.010	0.005	0.021	0.012	0.019	0.003	0.018
White%		0.018	0.029	-0.003	-0.011	-0.002	0.009	0.021	-0.027	-0.007	0.047	0.061	-0.017
Bachelor%		0.051	0.042	0.003	0.004	0.059	0.069	0.007	0.003	0.062	0.036	-0.013	0.008
Income		0.032	0.022	0.026	-0.002	0.015	0.041	0.044	-0.004	0.011	0.034	0.030	-0.000
Disability%		-0.003	-0.025	0.004	0.011	0.007	-0.009	0.019	0.015	0.006	-0.004	0.006	-0.008

Significant inequalities in PSM analysis are highlighted, where the green (red) color indicates consistent (contradictory) directions.

correlated with art & recreation choice in the New York MSA. After controlling for *white ratio*, the treatment effect of *bachelor ratio* on art choice is positive. Cases where the treatment effect reverses correlation direction are marked as bold in Table 2. These reversals of direction demonstrate the importance of PSM to identify inequalities from a causal perspective.

Counterfactual network embeddings

Justification of CRANE. We justify the effectiveness of CRANE by examining two stages, the counterfactual random walk stage and the network embedding learning stage, respectively.

First, we justify our proposed counterfactual random walk by examining its consistency with PSM. In the design of the counterfactual random walk, *observed outcomes* should characterize the characteristics of neighborhoods with higher urban facility accessibility. Therefore, for demographic features with significant positive treatment effects on urban facility accessibility given by PSM, *observed outcomes* are expected to be larger than the *alternative outcomes*. We count the total number of pairs among 77 demographic feature-POI category pairs that violate

this law on different sample sizes and draw their relationship in Fig. 5A. We find that when we sample more random walks, the violation percentage decreases rapidly and drops below 10% when the sample size reaches 100,000. Thus, with a sufficiently large sample size, the counterfactual random walk serves as an effective approximation of the traditional PSM procedure. In the following experiments, we sample 200,000 random walks for each category, ensuring the credibility of our analysis.

Second, we evaluate the effectiveness of our embedding algorithm by comparing each feature's treatment effect on urban resource access and the regression coefficient between $L(T)$ and its proximity with POI categories in the embedding space. We measure proximities with the inner product between $L(T)$ and POI category's embedding vectors. Regression coefficients are listed in Table 3 (See Supplementary Table S6 for significance levels). Insignificant treatment effects of demographic features on urban facility accessibility are marked in white cells. Among 77 significant demographic-category pairs ($p < 0.05$), 92.2% (71) pairs have identical directions in treatment effect with regression coefficient marked by green cells. 6 pairs that have the opposite

Table 4 Predicting the urban facility accessibility of out-of-sample neighborhoods.

Category	Input	NY	LA	Chi	Dal	Hou	DC
Art & Recreation	mobility statistics	14.13%	3.04%	9.72%	8.17%	- 2.50%	12.74%
	LINE embedding	19.58%	4.44%	15.70%	8.40%	1.63%	20.00%
	node2vec embedding	22.56%	4.72%	26.36%	7.68%	2.10%	18.84%
	CRANE embedding	37.18%	3.65%	33.66%	8.56%	2.43%	20.27%
Sports	mobility statistics	5.79%	0.57%	6.89%	5.16%	5.94%	3.35%
	LINE embedding	4.97%	1.01%	7.61%	6.66%	7.92%	4.56%
	node2vec embedding	4.71%	0.88%	7.12%	6.71%	8.17%	4.16%
	CRANE embedding	5.98%	1.06%	7.94%	7.29%	8.36%	5.03%
Education	mobility statistics	17.90%	7.79%	24.23%	8.58%	4.85%	10.54%
	LINE embedding	20.72%	10.66%	30.16%	9.20%	10.26%	7.79%
	node2vec embedding	23.49%	11.01%	28.15%	12.28%	11.25%	13.29%
	CRANE embedding	35.39%	11.95%	31.54%	14.99%	10.16%	14.27%
Health	mobility statistics	95.45%	277.39%	18.15%	27.07%	25.51%	39.90%
	LINE embedding	78.44%	292.94%	21.61%	15.83%	34.46%	33.15%
	node2vec embedding	100.61%	266.29%	19.39%	22.07%	36.78%	33.70%
	CRANE embedding	105.14%	276.85%	22.36%	33.84%	44.34%	32.09%

The values show the relative improvement in explained variance compared to using raw demographic features as model input. Bold text indicates the most improved method for prediction performance.

Table 5 Predicting the urban facility accessibility of the neighborhoods in Chicago MSA by transferring the knowledge from other MSAs.

Category	Input	NY	LA	Dal	Hou	DC
Art & Recreation	mobility statistics	8.31%	6.36%	2.89%	7.80%	4.92%
	LINE embedding	8.04%	11.79%	14.65%	11.12%	14.11%
	node2vec embedding	18.05%	26.98%	26.38%	21.50%	21.72%
	CRANE embedding	30.06%	30.26%	30.34%	26.52%	31.13%
Sports	mobility statistics	5.46%	5.28%	4.75%	5.33%	5.14%
	LINE embedding	6.81%	6.75%	6.92%	6.66%	6.73%
	node2vec embedding	6.93%	6.85%	7.17%	7.48%	6.86%
	CRANE embedding	7.80%	7.89%	7.71%	7.54%	7.61%
Education	mobility statistics	12.98%	13.75%	11.97%	12.81%	14.16%
	LINE embedding	29.31%	25.51%	31.35%	29.31%	33.07%
	node2vec embedding	28.41%	27.40%	31.65	29.30%	30.84%
	CRANE embedding	29.59%	30.00%	30.81%	29.41%	27.12%
Health	mobility statistics	10.90%	15.70%	12.42%	17.45%	15.09%
	LINE embedding	21.47%	20.46%	19.83%	21.16%	15.37%
	node2vec embedding	21.65%	18.62%	18.30%	23.49%	13.86%
	CRANE embedding	22.13%	22.04%	14.42%	11.43%	15.41%

The values show the relative improvement in explained variance compared to using raw demographic features as model input. Bold text indicates the most improved method for prediction performance.

direction are marked in red cells. This result confirms the consistency between the embedding algorithm and matching analysis.

We further illustrate the proximity between the POI category and demographic embeddings with the highest and lowest $L(T)$ in Chicago in Fig. 5B. Embedding vectors are projected to a 2D plane where proximities are preserved. To better demonstrate gaps in urban facility accessibility, we mark the direction from lowest $L(T)$ to highest $L(T)$ for each demographic feature in the plane. The visualization demonstrates (1) the association between high $L(bachelor\%)$ and high accessibility to sports and educational facilities, and (2) the association between high $L(white\%)$ and high accessibility to art and sports facilities. Based on the above observations, the counterfactual network representation learning method effectively captures MSA-level accessibility gaps across neighborhoods of different demographic backgrounds.

Predictive analysis. Given that the learned embedding vectors can successfully capture macro MSA-level treatment effects, we proceed to employ these vectors for predicting facility accessibility at the micro neighborhood-level, thus confirming their ability to

capture causal relations in fine-grained accessibility gaps. Specifically, we first conduct a within-city prediction task. In each MSA, we randomly select 80% of the neighborhoods as the training set, on which we perform *counterfactual random walks* and the network embedding algorithm. We combine the raw demographic features of each neighborhood with the dot products between their corresponding $L(T)$'s embedding vectors and all category embedding vectors as the input for Multilayer Perceptron (MLP) regression models to fit the neighborhood's accessibility to four categories from the training set. Then we test the fitted models on the remaining 20% neighborhoods and estimate performance as the average explained variance on five train/test random splits (see Supplementary Notes S1.3 for implementation details).

We compare CRANE to four baselines with different model input combinations. The first baseline only utilizes the neighborhood's raw demographic features. The second baseline utilizes both raw demographic features and aggregate statistics of urban facility accessibility from the training set. For instance, if a neighborhood has the highest $L(white\%)$, the average urban facility accessibilities of all neighborhoods with the highest

$L(\text{white}\%)$ in the training set are concatenated with this neighborhood's raw features as model input. The third and fourth baselines utilize embedding vectors learned by two graph embedding methods, namely LINE (Tang et al., 2015) and node2vec (Grover and Leskovec, 2016). The two embedding methods are correlation-based and do not incorporate specific designs to account for causal relationships. We construct a category-neighborhood-feature network with urban facility accessibility as edge weights and use unsupervised versions of these methods to learn embedding vectors that retain proximity. Dot products between treatment level vectors and all category vectors are combined with the neighborhood's demographic features as the input of regression models. The relative improvements of the three methods compared with raw features are listed in Table 4.

Among all the methods considered, CRANE outperforms others in 20 out of the 24 prediction tasks, particularly in the New York, Chicago, and Dallas datasets. On average, CRANE improves the explained variance in these tasks by 12.57%. This underscores the benefit of employing a causal matching strategy to learn embedding vectors for out-of-sample prediction tasks. Unlike correlation-based methods which are susceptible to confounding effects and distribution shifts between training and test sets, CRANE's counterfactual random walk strategy offers greater stability in prediction performance by capturing embedded micro-level causal relationships.

In addition to within-city prediction tasks, we further test CRANE's ability to transfer knowledge of micro-level accessibility gaps across MSAs. Specifically, we utilize embedding vectors and aggregate statistics learned from other MSA mobility networks to predict urban facility accessibility in Chicago MSA (see Supplementary Notes S1.3 for implementation details). Improvements in explained variance compared with raw features as input are listed in Table 5. Each column represents the performance of transferring knowledge from another MSA to Chicago. Our embedding methods achieve the best performance in 16 out of 20 tasks, especially in predicting art and sports choices. Consequently, CRANE effectively distills knowledge pertaining to micro-level accessibility gaps that are universal across MSAs, thereby enhancing its capacity to generate insights with limited data availability.

Visualizing maps of urban accessibility gap. Having demonstrated CRANE's effectiveness in capturing MSA- and neighborhood-level accessibility gaps, we now assess its capacity to provide micro POI-level insights through a case study in Houston. Suppose we want to query the top 10 sports POIs preferred by high-income neighborhoods. The first method is to directly look at the mobility statistics and retrieve the top 10 sports POIs with the highest accessibility from neighborhoods with the highest $L(\text{income})$. The second method, i.e., our CRANE method, is to retrieve the top 10 sports POIs whose embedding vectors have the highest inner product with the embedding vector of the highest $L(\text{income})$. As a comparison, we plot the spatial distribution of queried POIs in Fig. 6. We use black, red, and blue colors to denote POIs that are queried (1) by both methods, (2) only by mobility statistics, and (3) only by CRANE embeddings. Although all queried POIs are located within rich regions of the city, the two methods make largely different choices. The reason lies in a confounding variable, *white ratio*. In Houston, *average income* and *white ratio* are positively correlated and they are both positively correlated with accessibility to sports POIs. After distinguishing neighborhoods in Houston by the difference in their $L(\text{income})$ and $L(\text{white}\%)$, we observe that most sports POIs queried by mobility statistics (red pins) are located within regions where $L(\text{income})$ and $L(\text{white}\%)$ are both high (white color). In contrast, most sports POIs queried by embedding (blue pins) are

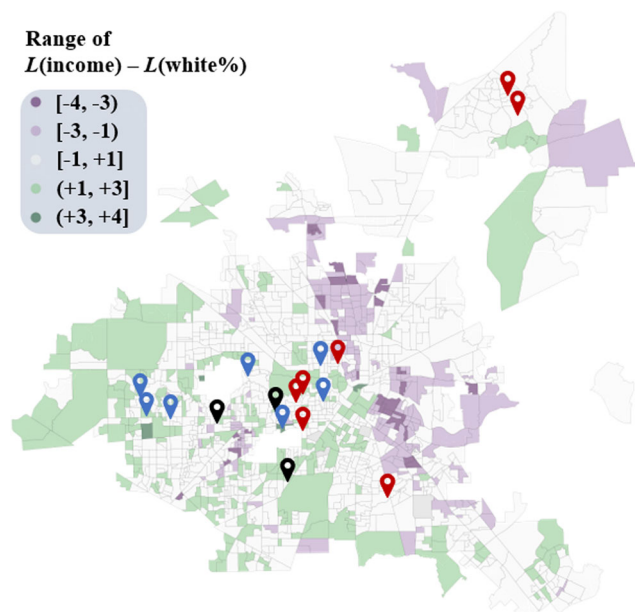


Fig. 6 The spatial distribution of queried POIs in Houston. We query sports POIs by two methods - retrieving POIs with the highest accessibility from high-income neighborhoods and POIs with the highest inner product with the embedding of the highest $L(\text{income})$. Red, blue, and black pointers represent POIs retrieved by the statistic method, embedding method, and both methods. Compared with red pointers, blue pointers locate in regions where the income level is greater than the white ratio level, capturing the advantage of income on accessing sports POIs in low white ratio neighborhoods.

located within regions where $L(\text{income})$ is greater than $L(\text{white}\%)$ (green color). This observation suggests that the advantage of high $L(\text{income})$ in sports accessibility among low $L(\text{white}\%)$ regions is captured and emphasized by CRANE, while unobserved in mobility statistics. Therefore, our CRANE framework can successfully untangle strong correlations between demographic features to generate POI-level causal representations.

Discussion

Traveling without limitation in the urban space is an essential capability for urban dwellers to sufficiently exploit urban opportunities. Our analysis of detailed mobility data covering two years reveals significant gaps in mobility behaviors from various demographic perspectives. Clear implications for urban planners and researchers can be drawn from the findings.

We first quantify the neighborhood-level treatment effects of demographic features on the mobility frequency before the COVID-19 pandemic by PSM. Significant treatment effects with identical direction among the majority of studied MSAs indicate serious inequality in mobility frequency. Neighborhoods with lower income have lower mobility frequencies. This is in line with long-standing social issues in class inequality across the U.S., which impairs equal sharing in urban opportunity. Such a gap might result from the interplay of the low-income group's segregation from the mainstream communities (Moro et al., 2021) and the high cost of urban transportation (Barbosa et al., 2021). In addition, the racial gap that neighborhoods with a more white population have higher mobility frequency is observed in all six MSAs but the capital city. Such class and racial inequalities in mobility frequency raise the concern on equality of citizens sharing opportunities in the urban space, calling for the design of urban policies that mitigate them. When we control race, income, and other confounding covariates, the bachelor ratio presents a

negative treatment effect in five MSAs. This evidence reflects the small commuting burden for highly-educated citizens under similar socioeconomic status. As opposed to the widely observed gender inequality in urban mobility in other countries (Gauvin et al., 2020; Ng and Acker, 2018), women's mobility frequency is not different from men's significantly in US cities. Additionally, we observe an insignificant treatment effect related to the disability household ratio, suggesting that urban spaces in the U.S. offer relatively equal opportunities for all genders and are accommodating for disabled individuals.

During the COVID-19 pandemic, mobility frequencies in large U.S. cities drop around 25% due to lockdown policies and POI shutdowns. By analyzing demographic features' effects on the reduction in mobility frequency in 2020, we reveal gaps in residents' capability of responding to an abrupt outbreak of infectious disease. Neighborhoods with a higher bachelor ratio reduce their mobility frequency more drastically in all six MSAs. Similarly, high-income neighborhoods also possess significant large mobility reductions in three MSAs. The above evidence reveals the gap in the capability to positively respond to mandate lockdown policies. Highly-educated people and people with better socioeconomic conditions have the advantage of accomplishing work remotely and living with their savings. By contrast, the nature of work for people with lower educational backgrounds requires them to travel in cities despite the social distancing and lockdown policies. The white population demonstrates better resilience during the pandemic shock. Neighborhoods with a higher white ratio have less mobility reduction in all six MSAs. This evidence suggests that the mobility gap between the white population and the minorities is exacerbated during the pandemic. The minorities are gradually pushed to the fringes of society, losing their rights and capability to exploit urban resources. Local government could assign economic support to the population in need and improve their remote working condition.

Furthermore, we analyze differences in residents' accessibility to different types of POIs. Both PSM and network representations demonstrate differences in accessibility to various urban venues. Sports venues are highly accessed by neighborhoods with higher average income, higher bachelor ratios, and higher white population ratios. Similarly, neighborhoods with higher average income and higher white population ratios have a higher proportion of visits to educational POIs. The class and racial gaps in education and sports accessibility urgently require policies striving for low-cost and quality opportunities for being educated and exercising for minorities and low-income populations. Neighborhoods with higher disability ratio and female ratio have higher urban facility accessibility to health POIs in New York, Chicago, Dallas, and Houston MSAs, indicating a high level of healthcare burden for women and people with disabilities.

Aside from the PSM method, we propose a counterfactual random walk-based representation learning algorithm called CRANE to capture the micro-mechanism of urban mobility inequality in an embedding space. Within- and cross-city prediction tasks confirm the effectiveness of the embedding vectors performing better and more stable in the prediction of urban facility accessibility. By querying proximity representation vectors in the embedding space, we can visualize the micro-level causal mechanism of gaps in urban facility accessibility. The effort of capturing accessibility gaps by representation learning approaches opens the door to understanding the micro-level mechanism of inequality and applying representation vectors for more urban tasks. For example, we can analyze POIs in which locations possess the harshest racial gaps. POI recommendations and site selection problems should also consider gaps in residents' mobility patterns in the embedding space to promote equal access to urban facilities.

It is possible that factors beyond our consideration may also contribute to gaps in mobility patterns, including spatial factors such as road network structures and the availability of public transportation. Our current focus is on assessing accessibility gaps across various demographic backgrounds. Given that demographic attributes largely determine residential choices and living contexts (Clark, 1992; Wilson, 2006), we have not included spatial factors as potential covariates that could confound demographic features. In future research, natural experiments could be conducted to analyze the impact of spatial factors on accessibility gaps. Here, by combining large-scale mobility data and leveraging causal analysis to balance key demographic covariates, we dramatically reduce confounding effects, providing an in-depth understanding of inequality in urban mobility. In addition, with the reopening policies and cancellation of mask mandates in years after 2021, consumption, production, and life are back on track. Future research can assess the inequality in the recovery of urban mobility in the post-pandemic era, revealing gaps in citizens' mobility resilience and how the pandemic shapes their mobility.

Overall, our study suggests that in urban planning and governance, more attention should be paid to ensuring equal accessibility to urban venues so that a city can truly become a sustainable place that "provides something to everybody" (Jacobs, 1961).

Data availability

The SafeGraph Monthly Patterns datasets that support the findings in this study are available from SafeGraph through the SafeGraph Data for Academics program. The availability of these data is under strict restriction under the Data License Agreement of SafeGraph. The data are used under the license for this study and are not publicly available. The conditions and limitations of access to the data can be found at <https://www.safegraph.com/academics>. The demographic data can be publicly obtained from U.S. Decennial Census and American Community Survey data at <https://www.safegraph.com/free-data/open-census-data>. Codes for reproducing the CRANE algorithm are available at <https://github.com/tsinghua-fib-lab/CRANE>.

Received: 13 June 2023; Accepted: 21 December 2023;
Published online: 09 January 2024

References

- Althoff T, Sosić R, Hicks JL, King AC, Delp SL, Leskovec J (2017) Large-scale physical activity data reveal worldwide activity inequality. *Nature* 547(7663):336–339
- Barbosa H, Hazarie S, Dickinson B, Bassolas A, Frank A, Kautz H, Sadilek A, Ramasco JJ, Ghoshal G (2021) Uncovering the socioeconomic facets of human mobility. *Sci Rep* 11(1):8616
- Bettencourt, LM. (2021) *Introduction to Urban Science: Evidence and Theory of Cities as Complex Systems*. MIT Press
- Carra G, Mulalic I, Fosgerau M, Barthelemy M (2016) Modelling the relation between income and commuting distance. *J R Soc Interface* 13(119):20160306
- Chang S, Pierson E, Koh PW, Gerardin J, Redbird B, Grusky D, Leskovec J (2021) Mobility network models of covid-19 explain inequities and inform reopening. *Nature* 589(7840):82–87
- Chen L, Xu F, Han Z, Tang K, Hui P, Evans J, Li Y (2022) Strategic covid-19 vaccine distribution can simultaneously elevate social utility and equity. *Nat Hum Behav* 6(11):1503–1514
- Clark WA (1992) Residential preferences and residential choices in a multiethnic context. *Demography* 29(3):451–466
- Dueñas M, Campi M, Olmos LE (2021) Changes in mobility and socioeconomic conditions during the covid-19 outbreak. *Hum Soc Sci Commun* 8(1):1–10
- Frias-Martinez V, Virseda-Jerez J, Frias-Martinez E (2012) On the relation between socio-economic status and physical mobility. *Inform Technol Dev* 18(2):91–106

- Garg N, Schiebinger L, Jurafsky D, Zou J (2018) Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proc Natl Acad Sci* 115(16):E3635–E3644
- Gauvin L, Tizzoni M, Piaggini S, Young A, Adler N, Verhulst S, Ferrer L, Cattuto C (2020) Gender gaps in urban mobility. *Hum Soc Sci Commun* 7(1):1–13
- Glaeser EL, Resseger M, and Tobio K (2008) Urban inequality. Harvard Institute of Economic Research Discussion Paper, (2168)
- Graells-Garrido E, Meta I, Serra-Burriel F, Reyes P, Cucchiatti FM (2020) Measuring spatial subdivisions in urban mobility with mobile phone data. In *Companion Proceedings of the Web Conference*. The Association for Computing Machinery, New York, p 485–494
- Griggs D, Stafford-Smith M, Gaffney O, Rockström J, Öhman MC, Shyamsundar P, Steffen W, Glaser G, Kanie N, Noble I (2013) Sustainable development goals for people and planet. *Nature* 495(7441):305–307
- Grover A, and Leskovec J (2016) node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. The Association for Computing Machinery, New York, p 855–864
- Hashtanasombat A, and Mascolo C (2019) Understanding the effects of the neighbourhood built environment on public health with open data. In *The World Wide Web Conference*. The Association for Computing Machinery, New York, p 648–658
- Jacobs J (1961) *The death and life of great American cities*. Random House, New York City
- Kadar C, Feuerriegel S, Noulas A, and Mascolo C (2020) Leveraging mobility flows from location technology platforms to test crime pattern theory in large cities. In *Proceedings of the international AAAI conference on web and social media*. The Association for the Advancement of Artificial Intelligence Press, volume 14, p 339–350
- Kingma DP, and Ba J (2015) Adam: A Method for Stochastic Optimization. In *International Conference on Learning Representations*, OpenReview.net, p 1–15
- Law R (1999) Beyond ‘women and transport’: towards new geographies of gender and daily mobility. *Prog Hum Geogr* 23(4):567–588
- Lenormand M, Louail T, Cantú-Ros OG, Picornell M, Herranz R, Arias JM, Barthelemy M, San Miguel M, Ramasco JJ (2015) Influence of sociodemographic characteristics on human mobility. *Sci Rep* 5(1):1–15
- Lund S, Madgavkar A, Manyika J, and Smit S (2020) What’s next for remote work: An analysis of 2,000 tasks, 800 jobs, and nine countries. McKinsey Global Institute, p 1–13
- Macedo M, Lotero L, Cardillo A, Menezes R, Barbosa H (2022) Differences in the spatial landscape of urban mobility: gender and socioeconomic perspectives. *Plos one* 17(3):e0260874
- McCullagh P (1980) Regression models for ordinal data. *J R Stat Soc: Series B* 42(2):109–127
- Miller HJ (2004) Tobler’s first law and spatial analysis. *Ann Assoc Am Geogr* 94(2):284–289
- 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
- Ng WS, and Acker A (2018) Understanding urban travel behaviour by gender for efficient and equitable transport policies. *International Transport Forum Discussion Paper*
- Pucher J, Renne JL (2003) Socioeconomics of urban travel: evidence from the 2001 nhts. *Transp Q* 57(3):49–77
- Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41–55
- Ryan J, Wretstrand A, Schmidt SM (2015) Exploring public transport as an element of older persons’ mobility: a capability approach perspective. *J Trans Geogr* 48:105–114
- Saha M, Chauhan D, Patil S, Kangas R, Heer J, Froehlich JE (2021) Urban accessibility as a socio-political problem: a multi-stakeholder analysis. *Proc ACM Hum-Comp Interact* 4(CSCW3):1–26
- Sallis JF, Conway TL, Cain KL, Carlson JA, Frank LD, Kerr J, Glanz K, Chapman JE, Saelens BE (2018) Neighborhood built environment and socioeconomic status in relation to physical activity, sedentary behavior, and weight status of adolescents. *Prevent Med* 110:47–54
- Saxon J (2021) Empirical measures of park use in american cities, and the demographic biases of spatial models. *Geogr Anal* 53(4):665–685
- Šćepanović S, Mishkovski I, Hui P, Nurminen JK, Ylä-Jääski A (2015) Mobile phone call data as a regional socio-economic proxy indicator. *PloS one* 10(4):e0124160
- Sen A (1980) *Equality of What? In McMurrin S Tanner Lectures on Human Values, Volume 1*. Cambridge University Press
- Sharkey P, Elwert F (2011) The legacy of disadvantage: Multigenerational neighborhood effects on cognitive ability. *Am J Soc* 116(6):1934–81
- Tang J, Qu M, Wang M, Zhang M, Yan J, and Mei Q (2015) Line: Large-scale information network embedding. In *Proceedings of the 24th international conference on world wide web*. The Association for Computing Machinery, New York, p 1067–1077
- Uteng, TP (2012) *Gender and Mobility in the Developing World*. World Bank, Washington DC
- Wang D, Pedreschi D, Song C, Giannotti F, and Barabasi AL (2011) Human mobility, social ties, and link prediction. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. The Association for Computing Machinery, New York, p 1100–1108
- Wang Q, Phillips NE, Small ML, Sampson RJ (2018) Urban mobility and neighborhood isolation in america’s 50 largest cities. *Proc Natl Acad Sci* 115(30):7735–7740
- Wilson WJ (2006) *The geography of opportunity: Race and housing choice in metropolitan America*. Rowman & Littlefield
- Xu Y, Belyi A, Bojic I, Ratti C (2018) Human mobility and socioeconomic status: analysis of singapore and boston. *Comput Environ Urban Syst* 72:51–67

Acknowledgements

This work was supported in part by the National Key Research and Development Program of China under grant 2020AAA0106000 to YL and the National Natural Science Foundation of China under U22B2057, 62272262 and U21B2036 to YL. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

YZ, FX, LC, YY, JE, LB and YL contributed ideas. YZ, FX, LC, YY and YL designed the research methods and provided the research outline. YZ performed the experiments and prepared the figures. JE, LB and YL provided critical revisions. All authors jointly participated in the writing of the manuscript.

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-023-02570-5>.

Correspondence and requests for materials should be addressed to Yong Li.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024