Revealing the relationship between human mobility and urban deprivation using geo-big data: a case study from London in the post-pandemic era

Tongxin Chen *[1], Xiaowei Gao †[1] and Tao Cheng ‡[1]

[1]SpaceTimeLab for Big Data Analytics, Department of Civil, Environmental and Geomatic Engineering, University College London, London WC1E 6BT, UK

GISRUK 2023

Summary

Investigating the association between human mobility and urban deprivation helps to understand the disparate routines of urban residents with different socioeconomic vulnerabilities. Though lots of research have revealed the difference in the population's mobility behaviours impacted by social distancing measures during the COVID-19 pandemic since 2020, limited analytics focuses on the inequality in mobility recovery patterns of urban residents in the post-pandemic era. Using a large-scale geo-big data set (mobile phone GPS trajectories), we calculated the associations between the measured mobility recovery rate and urban deprivation indices (seven categories) in 4835 London communities (LSOAs) during the first four months of 2022. We show that mobility recovery is associated with urban deprivation (particularly the 'Barriers to Housing and Services' deprivation index) over the observed post-pandemic period. The results further demonstrate that the residents from higher deprived/vulnerable communities are likely to obtain lower mobility recovery rates in London.

KEYWORDS: Human mobility, urban deprivation, vulnerability, geo big data, post-pandemic

*tongxin.chen.18@ucl.ac.uk
†xiaowei.gao.20@ucl.ac.uk
‡tao.cheng@ucl.ac.uk

1 Introduction

As non-pharmaceutical interventions to reduce the transmission of the COVID-19 virus, social distancing measures enacted by local governments in global cities have aroused rapid reduction in human mobility including population movement and activity since the pandemic outbreak (Jia et al., 2020; Flaxman et al., 2020; Hale et al., 2021). Though mobility restrictions were found efficient in containing virus infections in the urban population, such rapid social measures were identified to impose negative influences on societies including economic losses, fiscal revenue reduction and social inequality (Bonaccorsi et al., 2020; Lenzen et al., 2020). Several early COVID-19 studies have revealed the disparate responses in relation to mobility behaviours among urban populations with varied socioeconomic conditions, such as age, income, education and economic deprivation (Weill et al., 2020; Gauvin et al., 2021; Cheng et al., 2022). In particular, the socioeconomic inequality that the restrictions policies disproportionately hit the disadvantaged population groups also raises the discussion about the 'Luxury nature' or 'Privilege gap' during the ongoing pandemic (Huang et al., 2021; Goudeau et al., 2021).

With the pandemic controlled by the popularised vaccination, local public officials have continued to promote and calibrate the reopening plans to transition back to normality in global cities. For example, the UK government announced that all remaining legal COVID restrictions (e.g., passengers are also no longer required to wear face coverings on public transportation) were removed as parts of the 'Living with COVID plan' [1] on Feb 24, 2022. Limited research interpreting mobility and its relationship with social inequality among urban residents in the post-pandemic era. To address this concern, we first detected the mobility indicator and calculated the recovery rate by comparing it with the pre-pandemic level from a large-scale geo-big data set, then explored the association coefficient between mobility recovery rate with urban inequality represented by the deprivation index for each community (LSOA) in London during the post-pandemic (from Jan 1, 2022 to Apr 30, 2022). Our findings on relative mobility recovery associated with urban deprivation have important implications for policymakers to understand and manage inequality in cities as the prolonged effect of the COVID-19 pandemic.

2 Data and Methods

2.1 Mobile phone GPS data

In recent years, with the advancement of positioning technology, the proliferation of massive geo-tagged/-located big data sets reflecting dynamic human mobility/activity enables researchers to model human mobility within several urban studies, such as resilience (hazard management and vulnerability assessment) (Haraguchi et al., 2022), urban public health (Giles et al., 2020; Xiong et al., 2020), and urban crime (Levy et al., 2020; Chen et al., 2022). In this study, the human mobility geo-big data – mobile phone GPS trajectory data provided by Location Sciences AI [2] were collected from millions of mobile phone applications used by anonymous users under GDPR

[1] Prime Minister sets out plan for living with COVID: https://www.gov.uk/government/news/prime-minister-sets-out-plan-for-living-with-covid

[2] LSAI: https://www.locationsciences.ai/

compliance. The GPS data offer precise location coordinates from the Global Positioning System (GPS) and high-resolution timestamps of an individual's movement trajectory points.

2.2 Urban deprivation data

In this study, the latest urban deprivation indices (2019) of London were downloaded from the Ministry of Housing, Communities & Local Government website [3]. The 'Indices of Deprivation data' of London are the primary measurement of deprivation for 4835 small areas (community level), known as LSOAs. The main index is the 'Index of Multiple Deprivation (IMD)', which combines weighted measurement across seven distinct subtypes of aspects in deprivation including 'Income Score (rate)', 'Employment Score (rate)', 'Education, Skills and Training Score', 'Health Deprivation and Disability Score', 'Crime Score', 'Barriers to Housing and Services Score', and 'Living Environment Score'.

2.3 Human mobility measurement

2.3.1 Stay and home location detection

Stay detection. A *stay* is one user u spending a time period at one location, i.e., the GPS points are at/around the same location during an observed period (Zheng, 2015). So, for a user's GPS trajectory records/points $\mathbf{P}$, it can be denoted as a set of locations (coordinates) $\mathbf{l}$ with temporal information. So, each GPS point can be illustrated as $\mathbf{P}_i = (\mathbf{l}_i, t_i)$. As a stay trajectory $\mathbf{S}$ can be extracted/detected from $\mathbf{P}_i$, i.e., each stay can be denoted as $\mathbf{S}_i = \left(\mathbf{l}_i, t_i^{start}, t_i^{end}\right)$. Then, the stays can be extracted using the stay detection algorithm incorporating two parameters: Δd (the maximum Euclidean distance that the records around a location to define as a stay) and Δt (the minimum duration time that the records within the period to define as a stay). In this study, the Δd and Δt are defined as 5 mins and 50 metres (Chen et al., 2023), to delineate each user's stays from the GPS trajectory points for each day (120 days in our observation period).

Home location detection. In this study, the residents are defined as the users who obtain home locations in each day's routines. Then, we use a heuristic definition to detect individual home locations from the detected trajectory of stays, i.e., a user's (u) home location h is the stay location that the user visits the most frequently during the night-time period. In this study, we define the night-time period from 23 pm to 6 am to implement home location detection.

2.3.2 Mobility indicator and recovery rate

Radius of gyration. In this analysis, we calculate a classical human mobility indicator named radius of gyration (ROG) (i.e., a radial distance to a point) to characterise the typical distance travelled by a centre stay of a stay trajectory (Gonzalez et al., 2008). For a user's stay trajectory $\mathbf{S}$, the

[3] MHCLG: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019

radius of gyration is defined as:

$$rg(\mathbf{S};u) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(|\mathbf{S}_i,\mathbf{S}_m|)^2}. \tag{1}$$

For example, the radius of gyration in a user's stay trajectory (including four time-ordered stays) can be calculated as $rg = \sqrt{\frac{1}{5}(|\mathbf{S}_2,\mathbf{S}_0|)^2 + (|\mathbf{S}_2,\mathbf{S}_1|)^2 + (|\mathbf{S}_2,\mathbf{S}_3|)^2 + (|\mathbf{S}_2,\mathbf{S}_4|)^2}$. The $\mathbf{S}_0$ is the detected home location in a user's movement (i.e., stay trajectory) of a day.

Recovery rate. Once we complete the individual stay, home location detection and the mobility indicator calculation from each user's daily movement. Then, for one day, the ROG of each community (LSOA) is represented by the average ROG of residents (whose home location locates in the corresponding LSOA). Based upon the same procedures, the individual mobility indicator is calculated from the mobile phone GPS data set of Feb 2020 as the pre-pandemic period. So, the baseline ROG value is represented by the daily mean value of all LSOAs during Feb 2020 (29 days). At last, for each day and each LSOA, the ROG recovery rate is measured by the actual value divided by the baseline ROG value.

2.4 Models

To identify the association between relative mobility recovery and urban deprivation, we fitted the mobility recovery rate with eight deprivation indices (one main IMD and seven sub-types of deprivation indices) in 4835 LSOAs for each day (120 days in total) through a tree-based machine learning model named XGBoost which is widely recognised for high performance and interoperability (Mousa et al., 2018). In each fitting, the optimised parameters of the XGBoost regressor linking mobility recovery rate ($\mathbf{y}$) and eight deprivation indices ($\mathbf{X}$) are selected by minimising the root mean square error (RMSE) through the random search and 10-fold cross-validation [4] strategy. Then, we output coefficients of determinants (R^2) for representing the relationship between the mobility recovery and urban deprivation for each day. In addition, we also output the feature importance, which measures the value of the attributes in constructing the elevated trees in the trained XGBoost regressors. The feature importance is calculated by the weighted Gini purity values of all nodes within all decision trees in improving the number of performances by tree splitting (Chen and Guestrin, 2016).

3 Results

In the case study, we detected 280,000 residents (the user obtains home location) in London over the four months of 2022 and calculated the ROG recovery rate of all 4835 LSOAs during the 120-day observation period. For an overview of the human mobility recovery rate in London, figure 1 (a) illustrates the daily London ROG recovery rate, which is represented by the average ROG recovery rate of all LSOAs. The 7-day average trend shows the ROG recovery rate above baseline, mainly concentrated in the mid of February and April of 2022. In addition, the daily rhythms of the ROG

[4]scikit-learn: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.RandomizedSearchCV.html

recovery rate show distinct weekday-weekend patterns and the result of the Kruskal- Wallis H test (figure 1 (b)) denotes the statistically significant difference of ROG rates between the groups of weekends and weekdays. Figure 1 (c) demonstrates the spatial heterogeneity of the ROG recovery rates in London, where the highest ROG recovery rates are clustered in the city centre (each LSOA ROG recovery rate in the map is represented by the mean value of 120 days).

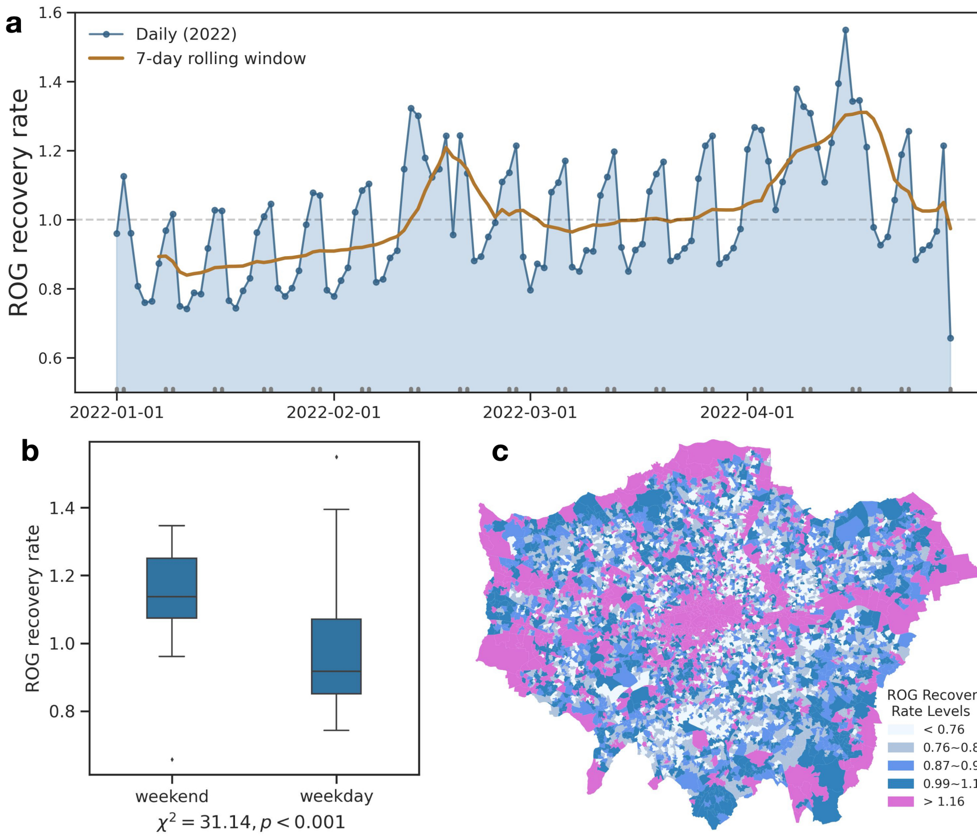

Figure 1: The overview of human mobility recovery rate in London. a) The daily variation of ROG recovery rate. b) The difference in ROG recovery rates between the groups of weekends and weekdays. c) The spatial distribution of ROG recovery rates.

After implementing the training process, we fitted the XGBoost regressors on the ROG recovery rates and derivation indices of 4835 LSOAs for each day. Figure 2 (a) denotes the distribution of the coefficients of determinants (R^2) of 120 fitted models. It is found that several coefficients over 0.5 show strong relationships between the response variables and the predictors in the corresponding models. Figure 2 (b) illustrates the mean feature importance of eight inputted variables of all 120 models. It shows the 'Barriers to Housing and Services Score' obtains the highest contribution and the 'Crime Score' obtains the lowest.

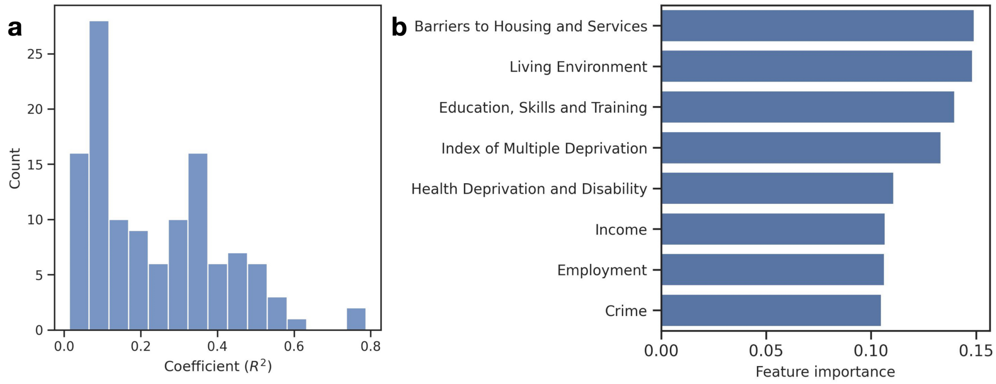

Figure 2: The association between mobility recovery and urban deprivation. a) The distribution of coefficients of determinants (R^2). b) The feature importance of XGBoost models.

4 Discussion and conclusions

Our analysis demonstrates the relationships between urban deprivation and human mobility during the post-pandemic. Specifically, here we explored how changes in mobility recovery rates were associated with working behaviours (the difference of ROG recovery rates between weekdays and weekends) and urban structure. Further, we find the highest feature importance of socioeconomic deprivation in discriminating the relative mobility change over the first four months of 2022. To discuss, Figure 3 shows the mobility recovery rates in three levels of deprivation ('Barriers to Housing and Services Score') in London (The daily ROG recovery rate of each deprivation group/level is represented by the mean value of all corresponding LSOAs). The 7-day average lines (bold ones) show a distinct segregation in ROG recovery rates in three deprivation levels with the highest deprived residents experiencing the lowest mobility.

The results facilitate our understanding of how urban deprivation influences mobility in different populations in London and highlight the residents from vulnerable communities who still experience the inequality of mobility recovery in the post-pandemic era. As the continual concerns in evaluating the mobility recovery in urban areas, further research is needed to understand the complex mechanisms that drive the relationship between mobility and socioeconomic deprivation and to examine the generalisability of the findings to other cities.

5 Acknowledgements

We acknowledge the anonymous reviewers for their valuable feedback and constructive criticism.

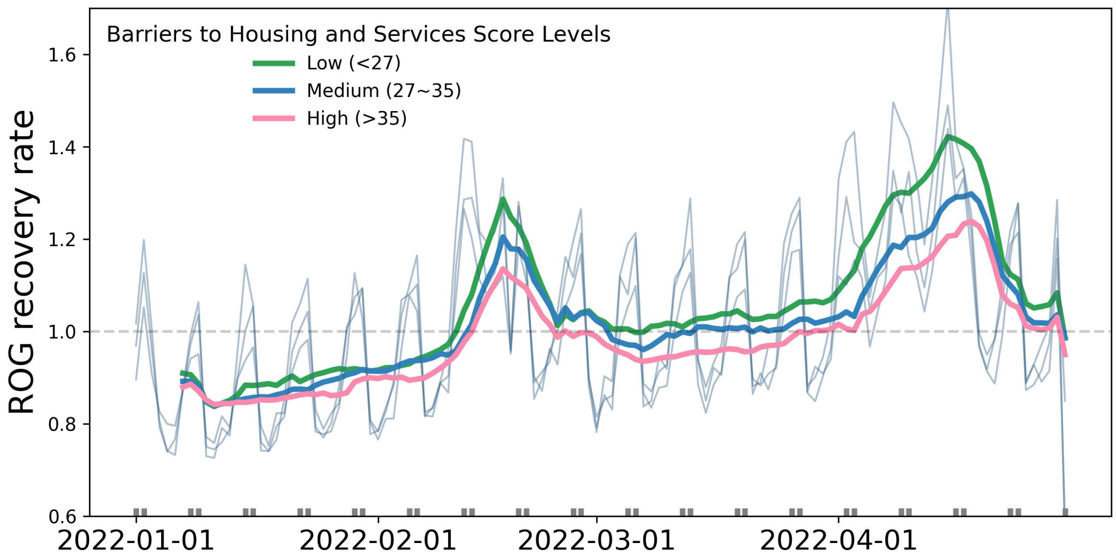

Figure 3: The mobility recovery rates in three levels of deprivation ('Barriers to Housing and Services Score') in London.

References

Bonaccorsi, G., Pierri, F., Cinelli, M., Flori, A., Galeazzi, A., Porcelli, F., Schmidt, A. L., Valensise, C. M., Scala, A., Quattrociocchi, W., et al. (2020). Economic and social consequences of human mobility restrictions under covid-19. *Proceedings of the National Academy of Sciences*, 117(27):15530–15535.

Chen, T., Bowers, K., Zhu, D., Gao, X., and Cheng, T. (2022). Spatio-temporal stratified associations between urban human activities and crime patterns: a case study in san francisco around the covid-19 stay-at-home mandate. *Computational Urban Science*, 2(1):1–12.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794.

Chen, T., Zhu, D., Cheng, T., Gao, X., and Chen, H. (2023). Sensing dynamic human activity zones using geo-tagged big data in greater london, uk during the covid-19 pandemic. *PloS one*, 18(1):e0277913.

Cheng, T., Chen, T., Liu, Y., Aldridge, R. W., Nguyen, V., Hayward, A. C., and Michie, S. (2022). Human mobility variations in response to restriction policies during the covid-19 pandemic: An analysis from the virus watch community cohort in england, uk. *Frontiers in public health*, 10.

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., Whittaker, C., Zhu, H., Berah, T., Eaton, J. W., et al. (2020). Estimating the effects of non-pharmaceutical interventions on covid-19 in europe. *Nature*, 584(7820):257–261.

Gauvin, L., Bajardi, P., Pepe, E., Lake, B., Privitera, F., and Tizzoni, M. (2021). Socio-economic determinants of mobility responses during the first wave of covid-19 in italy: from provinces to neighbourhoods. *Journal of The Royal Society Interface*, 18(181):20210092.

Giles, J. R., zu Erbach-Schoenberg, E., Tatem, A. J., Gardner, L., Bjørnstad, O. N., Metcalf, C., and Wesolowski, A. (2020). The duration of travel impacts the spatial dynamics of infectious diseases. *Proceedings of the National Academy of Sciences*, 117(36):22572–22579.

Gonzalez, M. C., Hidalgo, C. A., and Barabasi, A.-L. (2008). Understanding individual human mobility patterns. *nature*, 453(7196):779–782.

Goudeau, S., Sanrey, C., Stanczak, A., Manstead, A., and Darnon, C. (2021). Why lockdown and distance learning during the covid-19 pandemic are likely to increase the social class achievement gap. *Nature Human Behaviour*, 5(10):1273–1281.

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., et al. (2021). A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature human behaviour*, 5(4):529–538.

Haraguchi, M., Nishino, A., Kodaka, A., Allaire, M., Lall, U., Kuei-Hsien, L., Onda, K., Tsubouchi, K., and Kohtake, N. (2022). Human mobility data and analysis for urban resilience: A systematic review. *Environment and Planning B: Urban Analytics and City Science*, page 23998083221075634.

Huang, X., Li, Z., Jiang, Y., Ye, X., Deng, C., Zhang, J., and Li, X. (2021). The characteristics of multi-source mobility datasets and how they reveal the luxury nature of social distancing in the us during the covid-19 pandemic. *International Journal of Digital Earth*, 14(4):424–442.

Jia, J. S., Lu, X., Yuan, Y., Xu, G., Jia, J., and Christakis, N. A. (2020). Population flow drives spatio-temporal distribution of covid-19 in china. *Nature*, 582(7812):389–394.

Lenzen, M., Li, M., Malik, A., Pomponi, F., Sun, Y.-Y., Wiedmann, T., Faturay, F., Fry, J., Gallego, B., Geschke, A., et al. (2020). Global socio-economic losses and environmental gains from the coronavirus pandemic. *PloS one*, 15(7):e0235654.

Levy, B. L., Phillips, N. E., and Sampson, R. J. (2020). Triple disadvantage: Neighborhood networks of everyday urban mobility and violence in us cities. *American Sociological Review*, 85(6):925–956.

Mousa, S. R., Bakhit, P. R., Osman, O. A., and Ishak, S. (2018). A comparative analysis of tree-based ensemble methods for detecting imminent lane change maneuvers in connected vehicle environments. *Transportation Research Record*, 2672(42):268–279.

Weill, J. A., Stigler, M., Deschenes, O., and Springborn, M. R. (2020). Social distancing responses to covid-19 emergency declarations strongly differentiated by income. *Proceedings of the National Academy of Sciences*, 117(33):19658–19660.

Xiong, C., Hu, S., Yang, M., Luo, W., and Zhang, L. (2020). Mobile device data reveal the dynamics in a positive relationship between human mobility and covid-19 infections. *Proceedings of the National Academy of Sciences*, 117(44):27087–27089.

Zheng, Y. (2015). Trajectory data mining: an overview. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 6(3):1–41.

Biographies

Tongxin Chen, is a PhD candidate at SpaceTimeLab, Department of Civil, Environmental and Geomatic Engineering, University College London. His research interests include urban mobility, spatio-temporal big data analysis, crime science and applied machine learning.

Xiaowei Gao, is a PhD candidate at Spacetime Lab. His research interests are traffic mobility safety, Geo-AI methods and spatio-temporal data mining.

Tao Cheng, is a professor in Geoinformatics at the Department of Civil, Environmental and Geomatic Engineering, University College London. She is the Founder and Director of SpaceTimeLab for Big Data Analytics. Her research interests span network complexity, Geocomputation, space-time analytics and Big data mining (modelling, prediction, clustering, visualisation and simulation) with applications in transport, crime, business, health, social media, and natural hazards.