


Please cite the Published Version

Bannister, Jonathan  and O'Sullivan, Anthony (2021) Big Data in the City. Urban Studies: an international journal for research in urban studies, 58 (15). pp. 3061-3070. ISSN 0042-0980

DOI: <https://doi.org/10.1177/00420980211014124>

Publisher: SAGE Publications

Version: Published Version

Downloaded from: <https://e-space.mmu.ac.uk/627585/>

Usage rights:  [Creative Commons: Attribution 4.0](https://creativecommons.org/licenses/by/4.0/)

Additional Information: This is an Open Access article published in Urban Studies by SAGE Publications.

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

Big Data in the city

Jon Bannister 

Manchester Metropolitan University, UK

Anthony O'Sullivan

University of Glasgow, UK

Urban Studies

1–10

© Urban Studies Journal Limited 2021



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/00420980211014124

journals.sagepub.com/home/usj

Abstract

This editorial introduces a Special Issue on Big Data in the City. Collectively, six research articles and two commentaries explore the roles that Big Data can and might play in enhancing our understanding of urban processes and the qualities of urban outcomes. Big Data may be intrinsically considered a neutral technology but – refracted through existing power structures and resource distributions – its application within cities is by no means guaranteed always to help in the amelioration of social injustices or in the promotion of urban well-being. In application, Big Data becomes a performative technology that can be, is and will be further used in the creation and regulation of the cities of this century, a process that will be messy and of mixed consequence. The task for urban studies research is to shape that performativity, and to challenge any tendency that emerges to the further entrenchment of social inequities. In pursuit of these aims, and sensitively deployed, Big Data can be cast as part of the route map to better urban futures.

Keywords

Big Data, critical perspectives, performativity, urban condition, well-being

摘要

这篇社论介绍一期关于城市大数据的特刊。六篇研究文章和两篇评论共同探讨大数据在加强我们对城市过程和城市成果特征的理解方面可以发挥的作用。大数据在本质上可以被认为是一种中性技术，但通过现有的权力结构和资源分配折射出来的是，无法保证它在城市中的应用总是有助于改善社会不公正或促进城市福祉。在应用中，大数据成为一种可操演技术，可以、正在、并将进一步用于本世纪城市的创建和管理，这一过程将是混乱的，后果也是复杂的。城市研究的任务是塑造这种可操演性，并挑战任何进一步强化社会不平等的趋势。在追求这些目标的过程中，通过精心的部署，大数据可以成为通往更美好城市未来的路线图的一部分。

关键词

大数据、关键视角、可操演性、城市环境、福祉

Received March 2021; accepted March 2021

The rise of Big Data

Big Data, characterised by its volume, variety and velocity (Gandomi and Haider, 2015), is the new raw material of the 21st century. It has stimulated the evolution of multiple and diverse techniques (Gani et al., 2015) capable of mining and sifting this resource. Collectively, Big Data and Artificial Intelligence are serving to reshape the world and the way in which we understand it. Unsurprisingly, therefore, recent years have seen the launch of new journals such as *Big Data* (in 2013), *Big Data Research* and *Big Data and Society* (both in 2014), as well as stimulating multiple social science-relevant Big Data-themed special issues in other journals, such as: ‘Critiquing Big Data: Politics, Ethics, Epistemology’ (published in the *International Journal of Communication*, 2014); ‘Big Data Methods and Applications’ (published in the *Journal of Management Analytics*, 2015); ‘Big Data and Firm Performance’ (published in *British Journal of Management*, 2019); ‘Advances in Big Data Analytics and Intelligence’ (published in the *International Journal of Environmental Research and Public Health*, 2019); and ‘Big Data and the Human and Social Sciences’ (published in *Social Sciences*, 2021).

Whilst the explosion of interest in Big Data is entirely understandable (Burrows and Savage, 2014), to date *Urban Studies Journal* has gained (with notable exceptions) limited purchase in these debates. This Special Issue, we hope, will begin to rectify this shortfall. Its purpose is to draw together leading scholarship that explores and illustrates the potential of Big Data to illuminate the functioning of cities and the lived realities of their citizenry, and by doing so to enrich

urban studies. In these terms, the articles collated in this issue seek (as we sincerely hope that future submissions to the journal will) to illuminate some of the theoretical, empirical and methodological advantages and challenges for the city that rest in Big Data.

While the precise urban questions of actual or potential interest involving Big Data continue to multiply, this Special Issue has been shaped by two specific questions that we believe will always be of central concern to urban scholars. Firstly, how is Big Data being deployed in practice to enhance urban well-being? Secondly, how is Big Data being used to advance our understanding of the urban? In the remainder of this introduction, using these questions as frames of reference, we briefly sketch some of the contours of the emergent discourse on Big Data in the City, identifying also, and where appropriate, contributions to date from within *Urban Studies Journal*. We then introduce the novel contributions contained within this Special Issue.

Embracing the messy middle

Much of the debate on the implications of Big Data for urban well-being, while shaped by the inevitable uncertainty that surrounds substantive technological change, has been polarised around a restrictive binary centred on utopian and dystopian visions of the future (Boyd and Crawford, 2012). Thus, for cities and for urban living, Big Data and the smart technologies it feeds have been heralded as possessing the potential to redress urban maladies, to drive forward the functioning of cities and to improve the well-being of their citizenry (Kong and Woods, 2018; Shelton et al., 2015); they

Corresponding author:

Jon Bannister, Department of Sociology, Manchester Metropolitan University, Geoffrey Manton Building, Rosamond Street West, Manchester M15 6LL, UK.

Email: Jon.Bannister@mmu.ac.uk

have also been decried as ushering in an era of actual and potential intense surveillance and of widening inequalities (Curran and Smart, 2020; Lyon, 2014; Rieke et al., 2014). As ever, the truth is likely to be much more messy, multifaceted and complex. With respect to smart technology, surveillance practices deployed to support 'revanchist' responses to homelessness also introduce new capacity to facilitate supportive responses (Clarke and Parsell, 2019). Similarly with Big Data, a simplistic binary of hope and fear merely bookends emergent possibilities. A range of complex, nuanced, contradictory and tensioned outcomes, emanating from a simultaneous interplay of top-down attempts at technocratic order creation with the more chaotic bottom-up practices reflective of citizen urbanism (Barns, 2020), are actually likely to be the order of the day.

Whilst the current and future consequences of Big Data for societal well-being remain open to contestation, that it is having a profound transformative impact on urban environments is beyond dispute. Platform economies predicated on Big Data are increasingly serving to disrupt traditional forms of goods and services production, distribution and consumption. Exploring the case of Airbnb in London (UK), Ferreri and Sanyal (2018) show how sharing economy actors can thereby influence aspects of the governance of cities to suit corporate interests. Then again, perhaps the narrative of change is as powerful as its substance. Valdez et al. (2018), investigating the development of a smart transport application in Milton Keynes (UK), found identifiable benefits to derive from the reinforcement of existing city branding, through a smart city narrative that served to mobilise a network of actors behind the pursuit of smart region development, rather than from the technological and data-driven efficiency gains anticipated from the application itself. Beyond the

Global North, a more limited Big Data smart city narrative is unfolding within international urban studies discourse. Yet, there are notable exceptions. Chambers and Evans (2020), for example, consider how the Internet of Things (IoT) is being used to challenge the poor access to infrastructure and services experienced by populations living in informal settlements, noting that a substantial proportion of the global population reside in such settlements. In a case study of water and energy infrastructure in Nairobi, they show how IoT technology is being utilised to configure connections between users, providers and infrastructures.

Beyond its technological dimensions, Big Data involves issues of epistemology, ethics and social justice (Crawford et al., 2014). By posing challenges to the authority and value of the social sciences upon which it has traditionally rested, Big Data, by implication, poses significant challenges to urban studies per se. It has ushered in new actors, for example data scientists, and new forms of empiricism that threaten the death of theory, with the speed of the Big Data revolution seemingly outstripping the capacity of the social sciences to offer critical reflection (Boyd and Crawford, 2012; Burrows and Savage, 2014; Kitchin, 2014). Yet, even if the emerging research paradigm is likely to be characterised by statistical tools searching for increasingly sophisticated patterns in increasingly sophisticated data, these patterns will predominantly be urban patterns, the interpretation of which will require a disciplinary engagement that embraces the nature of urban reality. The need for an urban studies configured sociological imagination (Mills, 1959) will be no less pressing, and facts will continue to require examination through theoretical lenses to excavate meaning (Putnam, 2002).

The nature of urban studies inquiry, of course, has always been conditioned by the data and research technologies available,

which have varied greatly by time and place and continue to do so. There is nothing peculiar to urban studies in this. Yet, failure to recognise this conditionality does, potentially, hold major consequences, in terms of the misinterpretation of findings, the failure to capitalise on existing knowledge and the emergence of disciplinary fissure and fads. Social science in general has frequently been charged with faddishness, either in the topics chosen for investigation or in the methods used to investigate them (see *Economist*, 2016, for one such example). But what presents as fad in the deployment of novel techniques is often, rather prosaically, the beneficial outcome of overcoming constraints in both data availability and analytical capability.

Big Data, therefore, can facilitate exploration of areas of analytical interest already known, but previously unreachable, enabling theoretical as well as empirical advance (Mian and Rosenthal, 2016). It can, for example, illuminate the daily rhythms and activities of the city, creating 'rich databases of neighbourhood and other place-based contexts' (Sampson, 2013: 9). Qiang et al. (2020) provide one such example, in their examination of the urban population density function. This function, long thought to be the outcome of trade-off between housing price, commuting cost and employment, has previously not been possible to fully operationalise and test due to a lack of suitable commuting cost data. Leveraging crowd-sourced geospatial travel time data, these authors reassess population density functions for metropolitan statistical areas in the USA, contributing to a better understanding of urban morphology while also providing baseline information for monitoring and predicting future trends in urban population distribution conditioned by technological advance and environmental changes.

Extending this reasoning, Big Data, particularly when deployed in tandem with

traditional techniques and data, holds the potential to initiate substantive progress in urban studies (Sampson, 2019). Reades et al. (2018) exemplify this by taking advantage of recent developments in the field of machine learning to analyse socio-economic transition in London (UK) neighbourhoods and to predict those areas most likely to demonstrate future 'uplift' or 'decline'. They consider the implications of such modelling for the understanding of gentrification processes, noting that if qualitative work on gentrification and neighbourhood change is to offer more than a rigorous post-mortem, then intensive, qualitative case studies must be confronted with and complemented by predictions stemming from other, more extensive approaches. Howe (2021), in similar fashion, interweaves quantitative and qualitative data of people's everyday movements and the decision-making behind them, derived from volunteered geographic information from smartphones, to demonstrate the role that everyday movements play in driving urbanisation processes. Specifically, this macro- and micro-scale approach is used to highlight movement as a strategy for those living in poverty to access resources and subvert entrenched inequality.

As well as opening novel vistas, integrating big and small data also enables insight on the qualities of the data itself. Arribas-Bel and Bakens (2019) utilise Big Data collected from a location-based service in the Netherlands, to develop a rich catalogue of urban amenities and a measure of their popularity among users. By integrating this data with more traditional sources of socioeconomic data, the authors identify and quantify inherent biases in the Big Data resource, thereby establishing where it is likely to be useful and when it is going to be misleading. In a similar vein, Harten et al. (2021) utilise data scraped from the internet, in the form of classified advertisements, to examine Shanghai's hidden informal housing market, highlighting both the

possibilities and the pitfalls of using online content to study such informality. In the reporting process (Harten et al., 2021: 11), these authors sum up the situation succinctly: 'Big data is not always better data; it is different data.'

Big Data in the city

To be useful, Big Data requires adequate computational capacity, which in turn holds the prospect of creating as well as interpreting such data. Batty and Milton (2021) develop a web-based modelling framework capable under current technological constraints of running singly or in collaboration on smartphones and personal computers. Their framework allows the application of traditional land use–transportation interaction models to extensive spatial systems in real time, opening the way to enhanced scenario planning in support of more effective decision-making for the design and operation of better cities. A prototype model framework, called QUANT, is briefly demonstrated in application, using as exemplars the effects of a growth and decline in employment in a metropolitan area in North-west England, and of a new high speed subway line across London. The article innovates by offering the user an evaluative tool that supports repeated, near-instantaneous interrogation of spatial Big Data in real-world strategic land planning contexts. As Batty and Milton note, with the emergence of such modelling capacities, the effective future constraint on better strategic decision-making for cities becomes not data or computational power, but the education (and we would add incentivisation) of researchers, planners and policy makers in their application.

In relation to the distinction between 'big' and 'better', Bourassa et al. (2021) assess whether the inclusion of an employment accessibility index based on automobile

travel times collected from personal mobile devices can help improve hedonic price models used for residential property valuation. Using residential transactions data from Miami, FL (USA), this is achieved through a comparison of the Big Data index with another derived from a regional travel demand model. The Big Data approach is also assessed against distance-based measures of employment accessibility, geographic submarket representation and the use of regression models incorporating spatial lags on regressors and error terms. As Bourassa et al. point out, the differing conceptualisations align with differing theoretical interpretations of housing market operation. The main conclusions advanced are that the Big Data measure does not add meaningful explanatory or predictive power to hedonic models incorporating an employment accessibility index, and that models using geographic submarket dummy variables are of greater value than accessibility models, while a spatial autoregressive and spatial error approach outperforms other market representations. In this instance, the conclusion is that Big Data does not serve to improve the efficacy of hedonic modelling; more, and more granular, data do not equate to better data. As the authors note, this by no means undermines the potential of Big Data in housing market applications. Big Data may, over time, progress understanding of urban housing, through its potential for identifying the dimensions of population movement that are reflected in house prices. But much thought will be required to establish how best to achieve this.

While the demand for and supply of urban police services are co-constituted phenomena, they are typically analysed separately within criminological literatures, with lack of an adequate interplay dimension to previous research principally reflective of data limitations. Using artificial intelligence and multilevel modelling techniques, Ellison

et al. (2021) offer a new approach to determining the effectiveness, efficiency and fairness of urban area policing by combining Big Data on police deployment patterns and unstructured textual incident narratives across the large metropolitan region of Greater Manchester in the north-west of England with more traditional administrative data on calls for service. Their research makes use of Global Positioning System data to assess the resources consumed in frontline deployment to incidents, and the unstructured text narratives generated from received calls for service to assess some of the complexities embedded in each incident. Ellison et al. are able to demonstrate how policing demand and deployment associate across time and space with features of the urban environment, and how the place- and people-based complexities embedded in policing service calls shape the cumulative and marginal frontline resources expended in their address. Potential new insights into questions of public service value for money and policing legitimacy are thereby made possible.

Residence and race are well-addressed topics in the study of urban segregation. However, while interaction between racial groups depends upon where they travel during their everyday activities as much as it does on where they live, approaches to segregation that focus on across- rather than within-neighbourhood aspects of segregation are much thinner on the ground. Candipan et al. (2021) use Big Data to examine the nature of racial segregation in the contexts of everyday travel and neighbourhood connectedness. They propose a mobility-based measure they call the segregated mobility index (SMI) to capture the extent to which neighbourhoods by racial composition are connected to one another. Using geotagged tweets sent by Twitter users in the 50 largest cities in the USA, they find that segregated mobility patterns are

predicted by residential segregation and help to produce segregated urban neighbourhood networks, while the overall racial composition of cities and legacies of racial conflict also condition movement across neighbourhoods. The multidimensional and dynamic nature of segregation confirmed in this way also nicely illustrates the broader potential of Big Data for advancing understanding of the social organisation of cities.

Another illustration is provided by Wang and Vermeulen (2021), who explore the role of the built environment in maintaining neighbourhood vitality. Noting that urban design can both enhance and obstruct the potential for collective action, they use machine learning and computer vision algorithms to extract built environment features from images captured by Google Street View (GSV). The influence of these features (more specifically, the presence of car-related, walking-related and mixed-use land infrastructures) upon the survival rate of neighbourhood-based social organisations in Amsterdam, the Netherlands, is then explored using elastic net regression. In line with theoretical expectation, Wang and Vermeulen find that public and green spaces are positively associated with organisation survival rates, whilst the presence of environmental features that encourage car usage decreases organisation survival rates. Methodologically, the authors demonstrate the potential of Big Data and its associated technologies to identify and enumerate the fine detail of urban built environments objectively, quickly and, relative to survey approaches, cheaply. Substantively, the article points to new weapons for the armoury of urban planners and policy makers.

Liu and Miller (2021) show that the emergence of Big Data also creates potential for richer analytical and policy appreciations of, and greater service satisfactions from, urban public transportation systems. Their case study of the Central Ohio Transit Authority

in Columbus (USA) demonstrates that routinely collected real-time data can be used to develop measures for assessing the risk of, and consequent delays due to, missing bus transfers. Specifically, they interweave high-resolution schedule and real-time vehicle location data to create measures of Risk of Missing Transfers and Average Total Time Penalty. Using these measures, they simulate the potential of dedicated bus lanes to facilitate a reduction in both risk and delay. Ultimately, the measures generated by Liu and Miller, embedded in real-time applications, hold significant potential to inform urban individual travel decisions, as well as the operational and strategic decision-making of urban transit authorities.

The articles in this Special Issue by Batty and Milton, Bourassa et al., Ellison et al., Wang and Vermeulen, Candipan et al. and Liu and Miller help exemplify and demonstrate the value of Big Data in addressing both positive and normative urban issues, matters of 'is' and of 'ought', as well as how things work in practice and of how they can be improved. In an important contribution, Taylor (2021), however, reminds us that this alone is not enough. Big Data technologies applied within an urban systems research paradigm serve to re-present ageless epistemological and ontological problems in new liveries. What gets measured by Big Data, and what doesn't, is no neutral, unalterable fact of life. The light it can shine on some corners of urban life serves only to darken the shadows obscuring others. Improvements to urban living for some are often bought at invisible cost to other urban citizens, and Big Data-based policy prescriptions have as much power to inflict cost as they have to offer benefits. If we are blinded by the shininess of new research tools, we may fail to appreciate that we are applying them badly both as social scientists and as people. As Taylor says, high data granularity and volume do not guarantee a thick

description of urban systems, and the analytics performed on that data condition the realities of urban governance.

Barns (2021) further highlights the need for a constant sensitivity to the possible implications of Big Data for future urban outcomes, warning specifically against the danger of autonomous agents, created through Big Data, acting to replicate and reinforce existing social injustices in the urban sphere. Here we would have Big Data perpetuating an urban reality where reproduction rather than improvement has become the goal, and in which Big Data is simply the latest technology by which this is achieved. Urban studies research, Barns insists, must avoid becoming complicit in such a future, acting instead to challenge the systematic replication of 'unwanted routines', and to ensure that Big Data is instead put to use 'to create the kinds of cities worth replicating computationally'.

Moving forward

Big Data is, by its intrinsic nature, an urban phenomenon. It is rapidly establishing itself in some obvious areas of urban studies interest, notably governance, security and transportation, all often within the smart cities context, and the spatial impacts and differential effects on labour and product markets of platform urbanism. But the potential is yet barely scoped and the possibilities for broader application are manifold in areas ranging from planning to real estate, public order, segregation studies and neighbourhood vitality to name a few – and some of which are the subject matter of this Special Issue.

In publishing this collection, and by means thereof, we invite further Big Data contributions to *Urban Studies*, particularly with regard to new and under-represented areas of concern. Returning to where we began, we reaffirm a belief that the Big Data

contributions of greatest significance and lasting value will be those maintaining the clearest focus on using Big Data to advance our understanding of the urban condition and urban well-being. But in that context, we also affirm three caveats. Firstly, accepting the always-present temptation to over-apply new datasets, methods and technologies when they become available, sometimes in research contexts where that application is questionable at best and misleading at worst, it is worth re-emphasising that big is not necessarily better. Second, in exploring the betterment of society using Big Data, as with all other methods, a constant critical perspective on what betterment actually means remains essential; in assessing whether Big Data is improving the urban condition, who gets to say what is better, what is not and for whom is never to be considered a given. Finally, in using Big Data for positive analysis, it is important to remember that the city does not pre-exist as an eternal concept, static and unchanging, simply to be understood more deeply as new techniques and data allow. It is, rather, an inherently dynamic, in many respects performative, concept (Ashton et al., 2017; Shelton, 2017; Zook, 2017). For the good of the discipline, subtlety in the application of Big Data methods to urban studies must be matched with subtlety of treatment for the subjects to which it is being applied.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Jon Bannister  <https://orcid.org/0000-0002-1350-510X>

References

- Arribas-Bel D and Bakens J (2019) Use and validation of location-based services in urban research: An example with Dutch restaurants. *Urban Studies* 56(5): 868–884.
- Ashton P, Weber R and Zook M (2017) The cloud, the crowd, and the city: How new data practices reconfigure urban governance? *Big Data & Society* 4(1). DOI: 10.1177/2053951717706718.
- Barns S (2020) *Platform Urbanism: Negotiating Platform Ecosystems in Connected Cities*. Singapore: Palgrave Macmillan.
- Barns S (2021) Out of the loop? On the radical and the routine in urban big data. *Urban Studies* 58(15): 3203–3210.
- Batty M and Milton R (2021) A new framework for very large scale urban modelling. *Urban Studies* 58(15): 3071–3094.
- Bourassa SC, Hoesli M, Merlin L, et al. (2021) Big data, accessibility and urban house prices. *Urban Studies* 58(15): 3176–3195.
- Boyd D and Crawford K (2012) Critical questions for big data. *Information, Communication and Society* 15(5): 662–679.
- Burrows R and Savage M (2014) After the crisis? Big Data and the methodological challenges of empirical sociology. *Big Data & Society* 1(1). DOI: 10.1177/2053951714540280.
- Candipan J, Phillips NE, Sampson RJ, et al. (2021) From residence to movement: The nature of racial segregation in everyday urban mobility. *Urban Studies* 58(15): 3095–3117.
- Chambers J and Evans J (2020) Informal urbanism and the Internet of Things: Reliability, trust and the reconfiguration of infrastructure. *Urban Studies* 57(14): 2918–2935.
- Clarke A and Parsell C (2019) The potential for urban surveillance to help support people who are homeless: Evidence from Cairns, Australia. *Urban Studies* 56(10): 1951–1967.
- Crawford K, Miltner K and Gray M (2014) Critiquing Big Data: Politics, ethics, epistemology.

- International Journal of Communication* 8: 1663–1672.
- Curran D and Smart A (2020) Data-driven governance, smart urbanism and risk-class inequalities: Security and social credit in China. *Urban Studies* 58(3): 487–506.
- Economist* (2016) Economists are prone to fads, and the latest is machine learning. Available at: <https://www.economist.com/finance-and-economics/2016/11/24/economists-are-prone-to-fads-and-the-latest-is-machine-learning> (accessed 17 July 2017).
- Ellison M, Bannister J, Haleem M, et al. (2021) Understanding policing demand and deployment through the lens of the city and with the application of big data. *Urban Studies* 58(15): 3157–3175.
- Ferreri M and Sanyal R (2018) Platform economies and urban planning: Airbnb and regulated deregulation in London. *Urban Studies* 55(15): 3353–3368.
- Gandomi A and Haider M (2015) Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management* 35(2): 137–144.
- Gani A, Siddiqi A, Shamshirband S, et al. (2015) A survey on indexing techniques for big data: Taxonomy and performance evaluation. *Knowledge and Information Systems* 46(2): 241–284.
- Harten JG, Kim AM and Brazier C (2021) Real and fake data in Shanghai's informal rental housing market: Groundtruthing data scraped from the internet. *Urban Studies* 58(9): 1831–1845.
- Howe LB (2021) Thinking through people: The potential of volunteered geographic information for mobility and urban studies. *Urban Studies* 58(14): 3009–3028.
- Kitchin R (2014) Big Data, new epistemologies and paradigm shifts. *Big Data & Society* 1(1). DOI: 10.1177/2053951714528481.
- Kong L and Woods O (2018) The ideological alignment of smart urbanism in Singapore: Critical reflections on a political paradox. *Urban Studies* 55(4): 679–701.
- Liu L and Miller HJ (2021) Measuring risk of missing transfers in public transit systems using high-resolution schedule and real-time bus location data. *Urban Studies* 58(15): 3140–3156.
- Lyon D (2014) Surveillance, Snowden and big data: Capacities, consequences and critique. *Big Data & Society* 1(2). DOI: 10.1177/2053951714541861.
- Mian A and Rosenthal H (2016) Big Data in political economy. *The Russell Sage Foundation Journal of the Social Sciences* 2(7): 1–10.
- Mills CW (1959) *The Sociological Imagination*. New York, NY: Oxford University Press.
- Putnam H (2002) *The Collapse of the Fact/Value Dichotomy*. Cambridge, MA: Harvard University Press.
- Qiang Y, Xu J and Zhang G (2020) The shapes of US cities: Revisiting the classic population density functions using crowdsourced geospatial data. *Urban Studies* 57(10): 2147–2162.
- Reades J, De Souza J and Hubbard P (2018) Understanding urban gentrification through machine learning. *Urban Studies* 56(5): 922–942.
- Rieke A, Robinson D and Yu H (2014) Civil rights, Big Data and our algorithmic future. A September 2014 report on social justice and technology. *Upturn*. Available at: <https://bigdata.fairness.io/wp-content/uploads/2015/04/2015-04-20-Civil-Rights-Big-Data-and-Our-Algorithmic-Future-v1.2.pdf> (accessed 8 February 2021).
- Sampson R (2013) The place of context: A theory and strategy for criminology's hard problems. *Criminology* 51(1): 1–31.
- Sampson RJ (2019) Neighbourhood effects and beyond: Explaining the paradoxes of inequality in the changing American metropolis. *Urban Studies* 56(1): 3–32.
- Shelton T (2017) The urban geographical imagination in the age of Big Data. *Big Data & Society* 4(1). DOI: 10.1177/2053951716665129.
- Shelton T, Zook M and Wiig A (2015) The 'actually existing smart city'. *Cambridge Journal of Regions, Economy and Society* 8: 13–25.
- Taylor L (2021) The taming of chaos: Optimal cities and the state of the art in urban systems research. *Urban Studies* 58(15): 3196–3202.
- Valdez A-M, Cook M and Potter S (2018) Roadmaps to utopia: Tales of the smart city. *Urban Studies* 55(15): 3385–3403.

- Wang M and Vermeulen F (2021) Life between buildings from a street view image: What do big data analytics reveal about neighbourhood organisational vitality? *Urban Studies* 58(15): 3118–3139.
- Zook M (2017) Crowd-sourcing the smart city: Using big geosocial media metrics in urban governance. *Big Data & Society* 4(1). DOI: 10.1177/2053951717694384.