Sustainable Cities and Society 64 (2021) 102503

Contents lists available at ScienceDirect

Sustainable Cities and Society

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

Bias in smart city governance: How socio-spatial disparities in 311 complaint behavior impact the fairness of data-driven decisions

Constantine E. Kontokosta [a,b,*], Boyeong Hong [b]

[a] Marron Institute of Urban Management, New York University, New York, NY, United States
[b] Center for Urban Science and Progress, New York University, New York, NY, United States

ARTICLE INFO

Keywords:
Smart city governance
Data bias
Machine learning
Social equity
311
Complaint reporting

ABSTRACT

Governance and decision-making in "smart" cities increasingly rely on resident-reported data and data-driven methods to improve the efficiency of city operations and planning. However, the issue of bias in these data and the fairness of outcomes in smart cities has received relatively limited attention. This is a troubling and significant omission, as social equity should be a critical aspect of smart cities and needs to be addressed and accounted for in the use of new technologies and data tools. This paper examines bias in resident-reported data by analyzing socio-spatial disparities in '311' complaint behavior in Kansas City, Missouri. We utilize data from detailed 311 reports and a comprehensive resident satisfaction survey, and spatially join these data with code enforcement violations, neighborhood characteristics, and street condition assessments. We introduce a model to identify disparities in resident-government interactions and classify under- and over-reporting neighborhoods based on complaint behavior. Despite greater objective and subjective need, low-income and minority neighborhoods are less likely to report street condition or "nuisance" issues, while prioritizing more serious problems. Our findings form the basis for acknowledging and accounting for data bias in self-reported data, and contribute to the more equitable delivery of city services through bias-aware data-driven processes.

1. Introduction

The potential of urban data analytics and "smart city" technologies has been widely heralded as a means to improve the efficiency of city operations and quality of life (Bettencourt, 2014; Bouzguenda, Alalouch, & Fava, 2019; Duvier, Anand, & Oltean-Dumbrava, 2018; Glaeser, Kominers, Luca, & Naik, 2018; Horgan & Dimitrijević, 2019; Kontokosta, 2018; Wu, 2020). The rapidly growing urban data ecosystem – characterized by high dimensional, spatial-temporal data on everything from mobility patterns to household waste – has attracted researchers and practitioners to public sector applications of machine learning (Batty, 2012; Huang, Xie, Tay, & Wu, 2009; Kitchin, 2014; Kontokosta, 2018; Provost & Fawcett, 2013). Governance and decision-making in "smart" cities increasingly rely on resident-reported data and data-driven methods to support city management and planning (Bouzguenda et al., 2019; Horgan & Dimitrijević, 2019; Westraadt & Calitz, 2020). However, the issue of bias in these data and the fairness of outcomes in smart cities has received only limited attention in practice. This is a troubling and significant omission, as social equity should be a critical aspect of smart cities and needs to be addressed and accounted for in the use of new technologies and data tools. One particularly popular dataset used to train data-driven decision-making models is '311' complaints (Hartmann, Mainka, & Stock, 2017; Wu, 2020; Xu, Kwan, McLafferty, & Wang, 2017). More than one hundred North American cities, including New York City, Chicago, Toronto, Washington, DC, and Kansas City, use 311 systems to manage resident complaints and service requests and respond as needed (Kontokosta, Hong, & Korsberg, 2017; Layne & Lee, 2001; McClure, 2000; O'Brien, 2016a). As such, 311 provides a crucial link between residents and government and represents an example of co-production through digital technology (O'Brien, 2016a; Wu, 2020). Because resident reports provide a snapshot of conditions across a city in real-time, local governments are analyzing these data to understand and forecast problems, service demands, and quality-of-life issues, such as rodent infestations, illegally converted buildings, potholes, and heat and hot water outages (Johnson, 2010; Kontokosta, Hong, et al., 2017; Melkers & Thomas, 1998; O'Brien, 2016b; Schwester, Carrizales, & Holzer, 2009; Wang, Lingjing, & Sobolevsky, 2016).

However, people do not report local problems at the same rate; therefore, we presume that resident-reported data is not an objective

* Corresponding author at: 60 5th Avenue 2nd FL, New York, NY 10011, United States.
E-mail address: ckontokosta@nyu.edu (C.E. Kontokosta).

https://doi.org/10.1016/j.scs.2020.102503
Received 28 June 2020; Received in revised form 3 September 2020; Accepted 18 September 2020
Available online 8 October 2020
2210-6707/

representation of the actual conditions across a city. Some of this disparity may simply be a function of the problems individual communities are exposed to – a neighborhood with better conditions should elicit fewer complaints per person than one with poor conditions (Thijssen & Van Dooren, 2016). At the same time, two individuals facing similar conditions may have different responses based on their expectations for what conditions should be. For instance, an individual accustomed to seeing rodents in their building may be less likely to complain than someone seeing a rodent in their apartment for the first time. In addition, individuals may have varying levels of trust in government and differing expectations about whether the government will actually respond that make them more or less likely to report a problem (Christensen & Lægreid, 2005; Sjoberg, Mellon, & Peixoto, 2017; Teo, Srivastava, & Jiang, 2008; Welch, Hinnant, & Moon, 2004; Wu, 2020). Together, these factors influence individual levels of civic engagement that result in inconsistent patterns of reporting. Using these data for decision-making, then, can lead to biased outcomes – resulting from the over- or under-estimation of the location and severity of problematic conditions – that lead to unfair or inequitable city service delivery and an uneven response to actual problems (Barocas, Boyd, Friedler, & Wallach, 2017; Chouldechova, 2017; Drosou, Jagadish, Pitoura, & Stoyanovich, 2017; Kontokosta, Weiss, Snively, & Gulick, 2017; Schill & Wachter, 1995).

This paper examines whether, and to what extent, bias exists in resident-reported complaint data and evaluates the potential impact of observed bias on the fairness of data-driven urban decision-making models. We explore the disparities in the use of 311 and identify both the household and neighborhood factors that influence resident-government engagement. To do so, we analyze data from the City of Kansas City, Missouri (KCMO), which include a resident satisfaction survey of 21,046 responses from 2014 to 2017, more than 500,000 311 service reports, and actual (measured) street pavement condition assessments of 29,884 street segments. We analyze disparities in the use of 311 linked to the perception and the prioritization of local problems, focusing on physical street conditions. After comparing actual and expected geolocated complaint reporting on street conditions and observed street condition assessment scores, we classify under- and over-reporting neighborhoods and propose a bias-adjusted report weighting to account for observed reporting behaviors. As a validation of our approach, we analyze actual pothole repair services provided by KCMO in under- and over-reporting neighborhoods in order to identify inequities in city service delivery resulting from disparities in 311 reporting.

We overcome significant limitations in previous research on this topic by using geolocated complaint data and objective, "ground-truth" condition assessments to develop a new approach to quantitatively study data bias in city services. This allows us to move away from simple per capita complaint rates toward the measurement and understanding of complaint behavior in the context of the actual problems communities face. The goal is to develop a new tool to identify, assess, and account for bias in data-driven urban decision models trained on resident self-reported data. We find that neighborhoods characterized by low-income, minority populations are less likely to report street condition or "nuisance" issues, despite lower satisfaction with city services and lower street maintenance quality. The article is organized as follows: Section 2 begins with a discussion of relevant literature on public sector data-driven decision-making and the determinants of resident-government interactions and 311 reporting. Section 3t presents our data and methodology to identify, explain, and adjust for disparities in complaint behavior. Next, we present our findings and validate the models using observed condition assessment data, and then discuss the implications of our results and provide suggestions for accounting for bias in data-driven decisions.

2. Smart city governance and data-driven decision-making

The recent use of machine learning or artificial intelligence to enhance city service delivery has generated new concerns about fairness and bias. Many researchers, social justice advocates, and, more recently, policymakers, have begun to question the potential bias embedded in algorithmic decision-making (Garcia, 2016; Hajian, Bonchi, & Castillo, 2016; Zhang & Neill, 2016). While attention has been focused on bias in computer vision and facial recognition software, the application of machine learning methods to prioritize the allocation of limited city services and resources raises concerns given the "black box" nature of many of these tools (Ibrahim, Charlson, & Neill, 2020). The causes and consequences of algorithmic bias have attracted new research by data scientists, hoping to find technical solutions independent of the moral, ethical, and contested values that provide the context for defining and evaluating fairness and equity in social science (Amini, Soleimany, Schwarting, Bhatia, & Rus, 2019; Verma & Rubin, 2018). While accounting for bias within algorithms is an important step in data-driven decision-making, a significant challenge emerges from the data used to train these models. Data provenance – what is collected, when, how, by whom, and for what purpose – has historically been a cause of tension and debate when applied to quantitative analysis in urban policy and planning (Klosterman, 1994; Yiftachel, 1989). What is new in the current big data and machine intelligence era is the rapidly expanding collection, integration, and use of these data, in many cases by private sector firms, in ways that the general public often little understands.

One such data source is resident-generated complaint reporting through 311 platforms. The 311 system is one of the most widely used e-government systems, providing an important mechanism for resident-government interaction (Layne & Lee, 2001; McClure, 2000; Minkoff, 2016; O'Brien, 2016a). Residents can report problems and request non-emergency city services through a centralized platform, and reports are used to improve city agencies' allocation of resources (Johnson, 2010; Kontokosta, Hong, et al., 2017; Minkoff, 2016; O'Brien, 2016a). As such, 311 systems represent a collaborative model of co-production between resident-as-consumer and government-as-provider (Minkoff, 2016; O'Brien, 2016a; Thijssen & Van Dooren, 2016). Co-production processes significantly depend on the involvement and the engagement of all constituents for efficient and fair public service delivery (O'Brien, 2016a; Thijssen & Van Dooren, 2016; Thomas, 2013). However, the realization that not all residents equally participate in co-production processes has created equity concerns stemming from the representativeness of co-producers (Thijssen & Van Dooren, 2016). For instance, Cavallo, Lynch, and Scull (2014) focus on identifying those that do not use e-government systems. The authors analyze 311 service requests for parking meter repair, sidewalk and street conditions, potholes, and traffic signal issues in New York City, San Francisco, and Washington, DC at the census tract level. These conditions are selected based on the assumption that these types of problems are independent of socioeconomic status because road infrastructure is evenly distributed and available to all residents. The regression model for the New York City case indicates lower-income, elderly, female, African American, Hispanic, and households with children are less likely to report these problems. The authors suggest that local governments should target these neighborhoods to increase resident-government interactions and equal access to benefits. Levine and Gershenson (2014) use city service requests for snowplows through Boston's Constituent Relationship Management System (CRM), which is similar to other cities' 311 systems, to identify the relationship between rates of city service requests and ethnicity and immigration status. The authors find that certain racial groups and foreign-born residents tend to report less due to a lack of engagement in political processes. While Cavallo et al. (2014) and Levine and Gershenson (2014) focus on socioeconomic characteristics of residents to analyze the likelihood of reporting problems to local governments, O'Brien (2016a) seeks to understand how reporting propensity varies by territoriality of residents. The author uses 311 data for

Boston from 2010 to 2015 together with home address information of registered 311 users (representing 46% of all 311 users). By measuring distances between users' home locations and 311 report locations, the author finds that more than 80% of individuals report problems within 150 m from their homes. Also, the study highlights that reporting significantly increased with neighborhood-centric advertisement and engagement strategies. In a recent study of New York City 311 data, White and Trump (2018) find that relationships between contacting 311 and political participation vary depending on the type of political engagement, such as voting or political donations. This suggests that 311 data may not provide a simple proxy of civic participation.

Political participation, civic engagement, and "contacting" studies have focused on understanding the individual profile of co-producers or active participants of civic activities, including reporting crimes and engaging in public regulatory processes (Minkoff, 2016; O'Brien, 2016a; Thijssen & Van Dooren, 2016). Vedlitz, Dyer, and Durand (1980) use both linear and polynomial regression models to analyze the relationship between socioeconomic status and contacting rate to local governments in Detroit and Houston. The results indicate that local government contacting has a negative linear relationship with socioeconomic variables, such as age of housing, rent, and income (Vedlitz et al., 1980). In contrast, Jones, Greenberg, Kaufman, and Drew (1977) studied resident contacts and hypothesized a parabolic association between the rate of contacting and neighborhood social well-being, which the authors define in terms of housing age and distance to the central business district. The authors argue that while need for services declines with higher well-being, awareness of service delivery systems increases with well-being. The result is that neighborhoods in the middle range of well-being tend to report the most. An empirical analysis by Sharp (1984) of contacting in Kansas City finds a negative relationship between resident engagement and social well-being, reinforcing the work of Vedlitz et al. (1980). The rationale for this finding was that Kansas City had a centralized reporting system; therefore, it was assumed that awareness was constant across neighborhoods.

Hirlinger (1992) also focuses on identifying the relationship between resident-initiated reporting and the severity of problems, socioeconomic status, and political ties by using personal interview data. The results indicate that white households and young adults with past political activity are more likely to report issues (Hirlinger, 1992). A related study by Brady, Verba, and Schlozman (1995) attempts to understand the power of resources – time, money, and civic skills – to influence political participation. The authors highlight that resources represented as free time, higher income, and other organizational activities are significant predictors of political activity. The study emphasizes that different socioeconomic groups have access to varying level of resources, meaning that socioeconomic status has a systemic effect on civic participation (Brady et al., 1995).

Wu (2020) also examines variations in 311 reporting that stem from 311 users' demographic and socioeconomic characteristics. Using citizen survey data from San Francisco, the study finds a positive relationship between 311 use and technology acceptance and residents who frequently use public services, such as parks, libraries, or public transportation, are more likely to use 311. Conversely, the results indicate a negative relationship between citizen satisfaction and the frequent use of 311, reinforcing previous findings (Wu, 2020).

Many of the above studies exhibit non-trivial limitations to understanding actual reporting propensities and complaint behaviors across diverse socioeconomic and physical infrastructure contexts. Few, if any, of the existing 311 reporting studies account for variations in actual neighborhood problematic conditions, often due to the absence of ground-truth data and objective condition assessments. Examining resident reports without comparison to actual conditions creates an endogeneity problem, where one is attempting to unpack the who, when, and why of complaint reporting by using those same complaints as a measure of local need. In order to overcome these methodological challenges, we integrate complaint reporting with both perceived and actual measures of need, using objective condition assessments of street infrastructure to account for both the prevalence of problematic conditions and the differential response of residents to such problems. By accounting for subjective and objective measures of localized need, we are able to more fully understand reporting behavior as a function of neighborhood context. This allows us to quantify the disparate impact of decision-making processes that rely on 311 self-reported data and develop a bias-aware approach to accounting for observed bias in smart city decision tools.

3. Data and methods

3.1. Data

In order to identify disparities in 311 resident-government interactions, we use a range of datasets provided by DataKC (formerly the Kansas City Office of Performance Management) and the Kansas City, Missouri (KCMO) open data portal, as described in Table 1. Our primary data are Kansas City 311 (KC311) service requests, which include non-

Table 1
Data sources and descriptions.

Dataset	Time range	Granularity	Sample size	Source and description
311 service requests	2014–2017	Daily, GPS (X,Y)	400,036	All resident service requests and non-emergency complaints, provided by KC311 through OpenData KC
Street pavement score	2018	Street segment	29,884	Measurement of street condition, provided by DataKC
Citizen satisfaction survey	2014–2017	Quarterly, Address	21,046	Resident survey on level of satisfaction with city services, use of city services, and household characteristics each fiscal year provided by DataKC
Property violations	2014–2017	Daily, GPS (X,Y)	66,308	Neighborhood code enforcement violations issued by the KCMO Neighborhoods and Housing Services (NHS), provided by KCMO NHS through OpenData KC
Parcels	2017	Annual, Parcel	208,309	Shapefile of tax parcels, provided by OpenData KC
Crime	2014–2017	Minute, GPS (X,Y)	506,749	Crime location and type, provided by the KCMO Police Department through OpenData KC
Registered dangerous buildings	2014–2017	Parcel	1,438	Dangerous building cases evaluated in accordance with building code standards if they are a candidate for demolition, provided by KCMO NHS through OpenData KC
American Community Survey (ACS)	2016 5-year estimates	Annual, Census block group	4,506	Demographic and socioeconomic characteristics from the U.S. Decennial Census and American Community Survey

emergency complaints reported by residents about a wide range of local problems, such as property conditions, missed trash collection, animal control, street conditions, and parking issues. These data consist of the location of the reported problem, a timestamp, and a description of the request or complaint. Overall, the KC311 database consists of more than 1.33 million service requests since 2007, with 103,955 reports in 2016. Neighborhood and property issues are the most commonly reported (accounting for 53.6% [55,742] and 15.8% [16,434], respectively, of all service requests in 2016). Fig. 1 visualizes the spatial patterns of 311 reports in 2016 (showing a heat map, total volume, and population-normalized volume, respectively). DataKC provided the results of a resident satisfaction survey conducted between 2014 to 2017, consisting of 21,046 individual responses regarding resident satisfaction with, and use of, city services and infrastructure. The survey samples are stratified across city council districts based on a statistically significant random sample of the balanced population against census demographics (City of Kansas City, 2018). The format of the questionnaire is a five-point Likert scale with 5 being "very satisfied" and 1 being "very dissatisfied". We integrate and geolocate the KC311 data and survey responses with other KCMO data sources, including property maintenance code violations, designated "dangerous buildings" assigned by the KCMO Neighborhoods and Housing Services (NHS), crime data, and U.S. Census Bureau American Community Survey (ACS) neighborhood demographic data. To provide an objective measure of problematic conditions, we utilize a dataset of street condition assessment scores obtained through visual inspection by the KCMO Public Works Department Street Preservation program. We use these assessment scores, also known as the OCI (Overall Condition Index) rating, as a measure of actual street condition (objective need) for more than 29,000 individual street segments. The scores range from 0 to 100 based on the street pavement quality and need for repairs. The Kansas City Department of Public Works defines the OCI rating scale and the corresponding maintenance activities suggested for each rating as follows:

- 100–90: No treatment needed
- 80–90: Needs crack seal
- 65–80: Needs microsurfacing
- 40–65: Needs resurfacing (2 inches)
- 20–40: Needs rehabilitation (4 inches)
- 0–20: Needs reconstruction

For reference, a newly-paved street would receive a score of 100.

3.2. Methodology

Our hypothesis is that reporting rates will vary with the nature and extent of local problems and differences in household characteristics. Specifically, the use of 311 is a function not only of household socio-economic and demographic composition, but also of the relative severity of problems within that household's neighborhood. The question is not just a binary of whether a household reports an issue, but how do households prioritize complaint reporting when faced with a gradient of problematic conditions. Therefore, we analyze whether the propensity to contact 311 varies with neighborhood quality and level of service.

To do so, we focus on street conditions. More than 50% of Kansas City residents identified street infrastructure improvements as a priority for local government (KCStat, 2015). The Kansas City Department of Public Works is responsible for maintenance of over 2200 miles of roads within the City limits. As the street network is distributed throughout Kansas City, we would expect that the likelihood of reporting street-related problems should – in the absence of reporting bias – be proportionate to the number of households adjacent to a particular street segment, controlling for private vehicle use. We acknowledge that street complaints may be reported by non-residents, or those away from their home, which would not be identifiable in the 311 dataset, but consider the magnitude of non-resident reporting to be low. We estimate under- and over-reporting complaint rates by neighborhood by comparing the actual number of reported complaints and the expected number of complaints based on the observed street condition assessments.

Several data processing steps are needed to link individual street segment scores, reported complaints, and resident satisfaction levels. First, we create a fishnet map of 500 m × 500 m (0.3 mile × 0.3 mile) rectangular cells to aggregate data with different spatial resolutions (e.g. point location versus census block group) to the same geographical unit, as illustrated in Fig. 2. The grid cells are considered proxies for neighborhoods, as we assume that residents contact 311 to report local problems proximate to their place of residence (O'Brien, 2016a).

In order to develop a neighborhood classification matrix of 311 reporting propensities, we then estimate the expected street condition complaint volume for each of the 29,000 street segments. Our assumption is that street condition complaint volume is a function of (1) actual street condition ratings, (2) length of street segments, and (3) the number of adjacent households along a given street segment. Zha and Veloso (2014) demonstrate that most complaint types are positively correlated with population size, and we account for this by normalizing

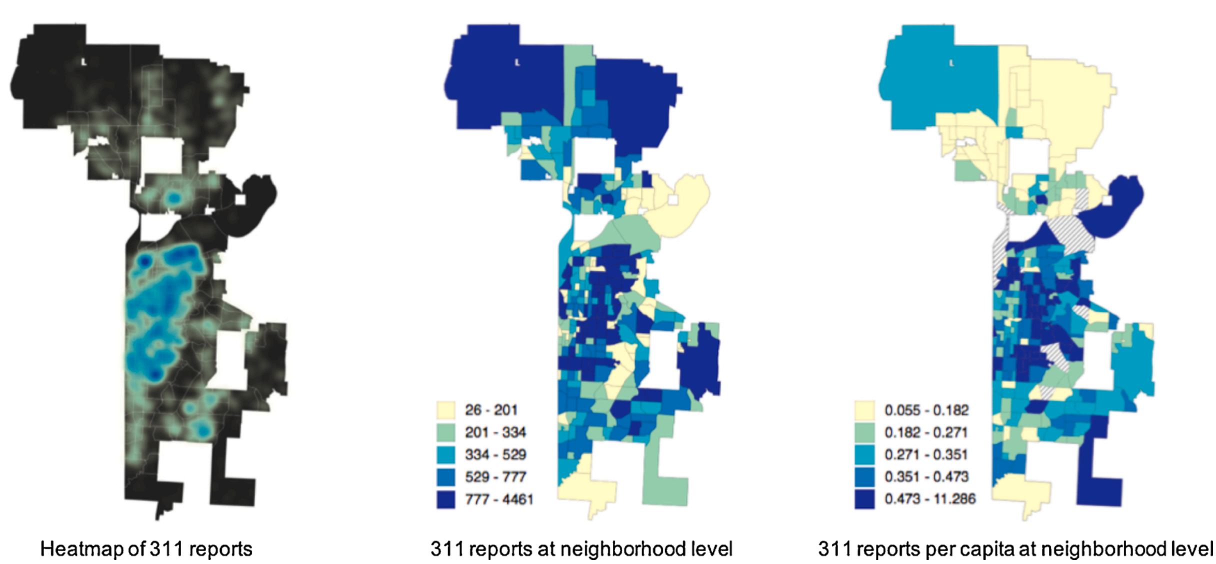

Fig. 1. Spatial patterns of 311 reports in 2016 (total 103,955, neighborhood level for the maps at center and right. Hatched areas indicate no residential population).

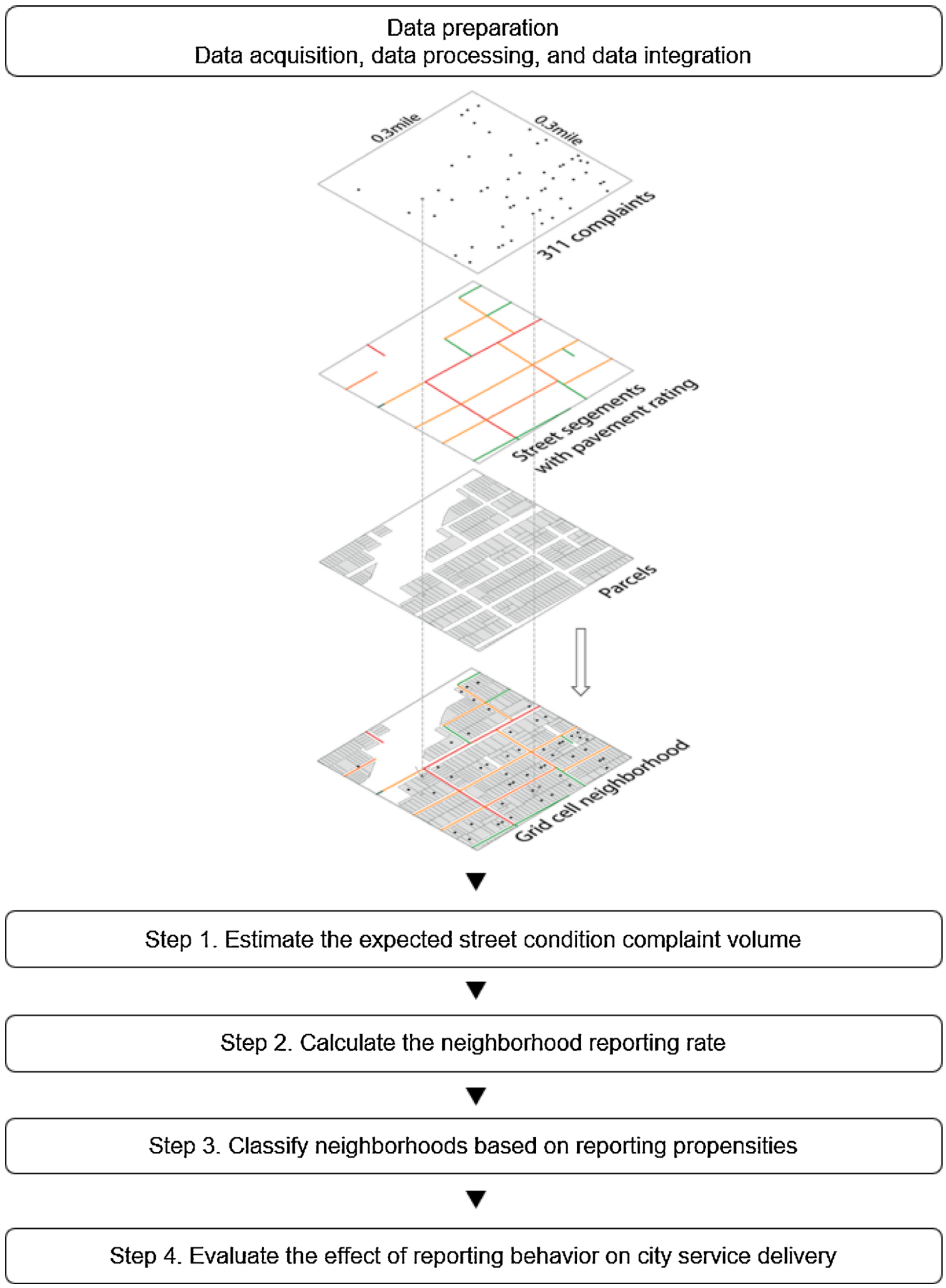

Fig. 2. Methodological approach and data integration schematic – parcels, street segments, 311 complaints, and other data are aggregated to 500 m × 500 m (0.3 mile × 0.3 mile) grid cells.

by the number of households living adjacent to a given street segment. As 311 calls are represent a "need for services" (White & Trump, 2018), we assume that households adjacent to streets with poor condition ratings are more likely to report a complaint. Based on this assumption, the 47,952 street condition complaints during the period from 2014 to 2017 are assigned to each neighborhood grid cell using the formula:

$$C_{\text{total}} = \sum C_k L_k H_k \tag{1}$$

where C_k is the number of complaints normalized per street mile per parcel, L_k is the total length of street segments within neighborhood k, and H_k is number of parcels in neighborhood k. C_k is given by:

$$C_k = P_k u, \forall k \in K \tag{2}$$

where P_k is the probability of complaint reporting based on the street condition ratings (adjusted to a 0 to 1 scale) for neighborhood k, and u is the citywide average number of street condition complaints per street mile per parcel ($u = 0.0036$). Specifically, P_k is calculated by:

$$P_k = \frac{(100 - S_k)}{100} \tag{3}$$

where S_k is the average street pavement scores of street segments within neighborhood k, scaled from 0 to 100. After computing the expected complaint volume for each neighborhood, the neighborhood reporting rate for street condition complaints is calculated as:

$$R_k = \frac{\mathrm{AC}_k}{\mathrm{EC}_k} \tag{4}$$

where R_k is the estimated reporting rate for neighborhood k, AC_k is the actual number of street condition 311 complaints from neighborhood k, and EC_k is the expected number of complaints based on Eq. (1). We then classify neighborhoods based on the computed reporting rates as under-reporting (estimated reporting rate lower than 1) and over-reporting (estimated reporting rate higher than 1).

If street conditions for two street segments in different neighborhoods are the same, how are resources allocated between those neighborhoods? Using 311-based service delivery prioritization, the neighborhood that reports the most will receive services more often, and thus this approach is sensitive to disparities in reporting propensities. In order to evaluate the effect of reporting behavior on city service delivery, we examine the widespread problem, in Kansas City and elsewhere, of potholes. Pothole repairs represent a significant complaint of KCMO residents and account for approximately $17 million per year in city budget allocation through the Street Preservation Program. As importantly for this analysis, potholes should be equally distributed across the City's street infrastructure, controlling for repaving or other major roadwork. Therefore, as an objective outcome measure of service delivery, potholes represent a unique opportunity to evaluate the fairness of various data-driven mechanisms to allocate city resources across different neighborhoods.

To process the necessary pothole complaint and repair data, we first use Python web scraping (BeautifulSoup libaray) and text processing to extract detailed information on pothole complaints and outcomes from the 2017 KC311 pothole case history *.html* pages (a total of 2,487). Specifically, we collect: (1) report date, (2) closed date, (3) a flag for duplicated cases from multiple reports, (4) whether a case is out of the City's responsibilities, and (5) whether a case is physically fixed (filled, repaired, or resurfaced). Based on these data, for each neighborhood we assess the level of pothole repair services based on the total number of repairs and mean time to resolution. We then compare these outcomes to the actual street condition scores and the resident satisfaction responses for under- and over- reporting neighborhoods using three different resource allocation methods: 311 complaints, street condition scores, and our proposed bias-adjusted 311 approach. In the absence of reporting bias, there should be no difference in the extent of street repairs between the over- and under-reporting groups, controlling for need.

4. Results: Evidence of disparities in reporting behavior

4.1. Neighborhood under- and over-reporting

Fig. 3 shows reporting rates for street condition complaints based on expected and actual complaint volume. Red points are considered over-reporting neighborhoods (reporting rate is higher than 1.0), while blue represents the under-reporting group. The black arrows illustrate the magnitude of over- or under-reporting of example neighborhoods based on the identity line. These results are spatially represented in Fig. 4 as maps of reported complaints, expected complaints, and under- and over-reporting neighborhoods. The 583 neighborhood grid cells in red are classified as over-reporting, while the 407 neighborhood grid cells in blue are under-reporting. The average street pavement scores for the over- and under-reporting groups are 68.27 and 48.00, respectively. This 20-point difference in street score is equivalent to two years of street deterioration based on Kansas City Department of Public Works maintenance standards. Moreover, the average reported resident satisfaction level for street maintenance (a measure of perceived need) for the over-reporting neighborhoods is 2.91, while the under-reporting group is significantly lower, at 2.15. Despite expressed dissatisfaction with street maintenance and observed poor street conditions (both objective need and perceived need are high), residents in the under-reporting neighborhoods do not report street condition complaints at a rate one would expect from conditions alone.

To explore socio-spatial variations in reporting behavior, household and neighborhood characteristics of the under- and over-reporting

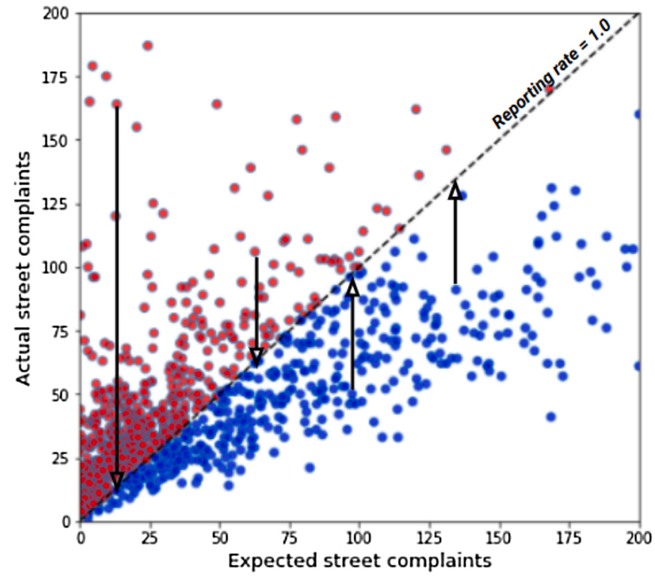

Fig. 3. Scatter plot of expected and actual street condition complaints – the black dashed line represents the reporting rate threshold. Red points are neighborhoods in the over-reporting group, while blue points are in the under-reporting group.

neighborhoods are presented in Table 2. Under-reporting neighborhoods are shown to consist of predominantly lower-income, minority households living in sub-standard housing conditions, with significantly higher proportions of vacant and dangerous properties and lower street condition scores. On the other hand, over-reporting neighborhoods are comprised of higher income, non-Hispanic white households with better overall neighborhood quality. These neighborhoods tend to contact 311 about street maintenance at the highest rates, despite high overall satisfaction with government services and high average street condition scores.

In addition to street conditions, over-reporting neighborhoods most frequently complain about "nuisance" issues, including sidewalks, trash, and trees. Collectively, these account for more than 51% of all complaints. In contrast, neighborhoods classified as under-reporting primarily report public safety issues, particularly property conditions. Therefore, the problem of representativeness in 311 reporting is not simply an issue of *how much* different individuals or neighborhoods report, but *the severity* of problems they experience and choose to report.

Figs. 5 and 6 further illustrate the relationships between neighborhood median income, service need, and contacting propensity for street condition and property condition problems. The need measure for public infrastructure is the inverse value of the street condition score, while the need measure for property condition is the inverse value of the resident satisfaction survey rating for code enforcement. Contact propensity is measured as the percentage of complaints reported for that issue. Regardless of the nature of the problem, there are strong negative associations between need and income, the result of areas with higher income experiencing fewer problematic conditions and better overall neighborhood quality. While previous work (Sharp, 1984) theorizes that contacting propensity should be negatively correlated with social well-being as a function of need for service, our analysis finds that contacting propensity patterns differ depending on the nature of the problem analyzed. There is a positive association between contacting propensity and income for public infrastructure issues (Fig. 5), while contacting propensity for code enforcement issues is negatively associated with income (Fig. 6). These distinctive patterns reinforce the finding that reporting behavior is a function of neighborhood context, problem severity, and individuals' perception and prioritization of local problems. Our results indicate that poor neighborhood quality forces

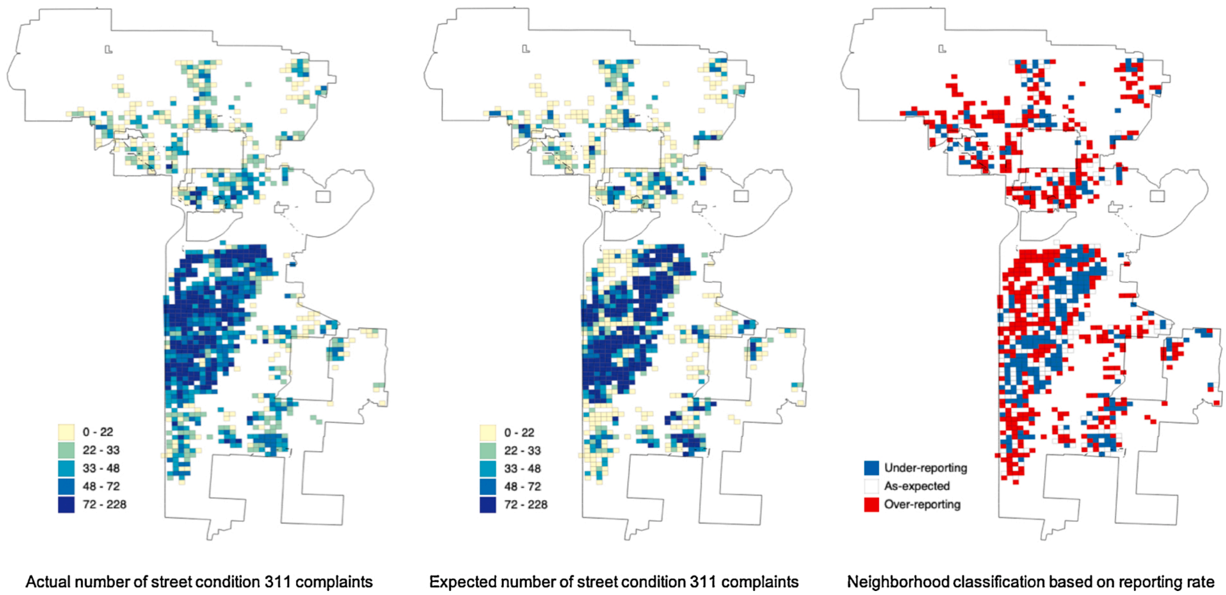

Fig. 4. Neighborhood classification based on reporting propensity.

Table 2
Demographic, socioeconomic, and neighborhood characteristics of under- and over- reporting neighborhoods.

Characteristics	Under-reporting	Over-reporting
Demographic and socioeconomic		
White	33.69%	80.25%
Black	54.67%	10.81%
Other races	9.66%	5.74%
Hispanic	11.61%	6.37%
Education attainment: high school degree	30.11%	18.04%
Education attainment: graduate school degree	6.09%	15.72%
Households in poverty	25.77%	7.86%
Income (median)	\$35,626	\$76,768
Employment rate	65.51%	71.92%
Household size (mean)	2.45	2.56
Household size – homeowner (mean)	2.44	2.70
Household size – renter (mean)	2.50	2.48
No. of rooms (median)	5.55	6.21
Year built (median)	1956	1984
Neighborhood		
Vacant parcels	23.28	4.84
Dangerous buildings	3.21	0.09
No. of crimes (mean)	568.96	115.01
Street pavement score (mean)	43.00	68.27

residents to prioritize 311 complaints on the most serious issues, and therefore there is a tendency to under-report quality-of-life problems, such as street conditions.

5. Pothole repairs: A case study

Table 3 summarizes pothole repair activity in under- and over-reporting neighborhoods. We observe significant differences between the two groups with respect to 311 pothole complaints and resulting outcomes. First, under-reporting neighborhoods reported 0.0056 pothole complaints per street mile per parcel in 2017 compared to 0.058 in over-reporting neighborhoods, an order of magnitude difference. Over-reporting neighborhoods filed more than one complaint for an individual pothole approximately 19% of the time, compared to less than 4% for under-reporting neighborhoods. As importantly, the number of 311 complaints per *fixed* pothole for the over-reporting group was 0.58, which is 1.53 times more than that in the under-reporting neighborhoods (0.38). In total, 958 potholes were repaired in over-reporting neighborhoods and 644 in under-reporting areas, representing normalized values per street mile per parcel of 2.91 and 2.40, respectively. The median time to resolution (from complaint to repair) for the two groups is the same at 4 days, but this should be considered in context since under-reporting neighborhoods are more likely to have lower street quality.

In order to provide operational insights for data-driven city service delivery, we apply our bias-adjusted 311 complaint metric (represented in Fig. 3) to pothole complaints. The number of pothole complaints reported in a given neighborhood is weighted by the expected complaint rate given neighborhood characteristics and controlling for problematic street conditions. In other words, the ratio of EC_k to AC_k from Eq. (4) is applied to actual 311 pothole complaints to estimate bias-adjusted pothole complaint volume. As previously described, the average number of pothole complaints reported in under-reporting neighborhoods is significantly lower than that in the over-reporting group. However, after adjusting complaint volume, there is no statistical difference between the over- and under-reporting groups with respect to normalized pothole complaints per street mile per parcel. The resultant average neighborhood adjusted complaint volume in the under-reporting neighborhoods is higher, at 3.90, than that in the over-reporting group (1.61).

How pothole repair services are prioritized across the city has clear implications for the equitable allocation of limited public resources. We have shown that both perceived and actual need for street infrastructure repairs are higher in lower-income neighborhoods. Yet these neighborhoods are less likely to report such problems, which may result in the under-allocation of city services in the places that need them most. To explore this further, we test three methods for pothole repair allocation across the city: based on 311 complaints only, based on visual inspection of street condition (OCI score) only, and based on our bias-adjusted 311 complaint metric. In the first method, the expected number of pothole repairs is proportional to the number of 311 pothole complaints regardless of street conditions or neighborhood reporting behaviors. The second method is based only on street pavement scores; therefore, a neighborhood with poorer street condition is assigned more pothole repairs, based on an expected number of potholes per mile for a given street score. Lastly, the bias-adjusted 311 complaint metric is used to allocate pothole repairs to neighborhoods taking into account both 311 reporting propensities and actual street conditions. A summary of the results is shown in Table 4.

A clear pattern emerges in the distribution of pothole repairs across

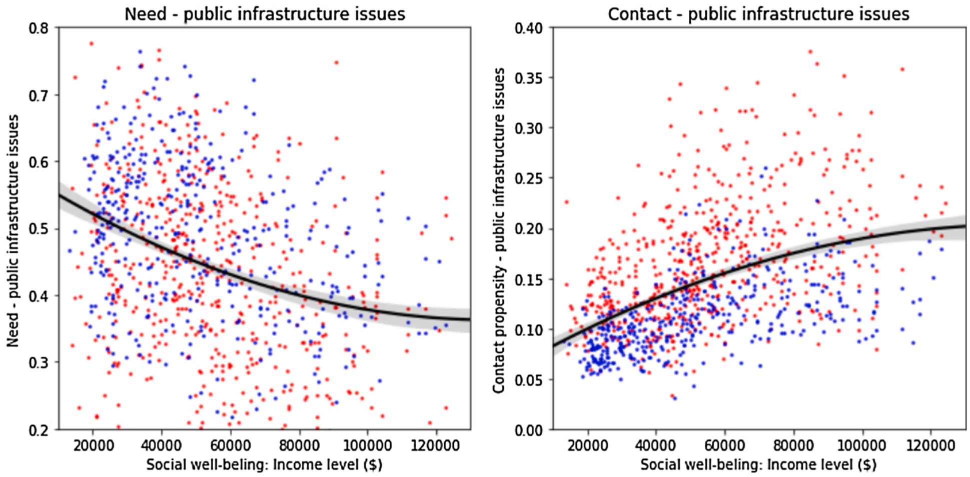

Fig. 5. Need and contacting propensity – public infrastructure. (The need measure for public infrastructure is the inverse value of actual street condition score. Colors are based on the classification results described in Fig. 4.)

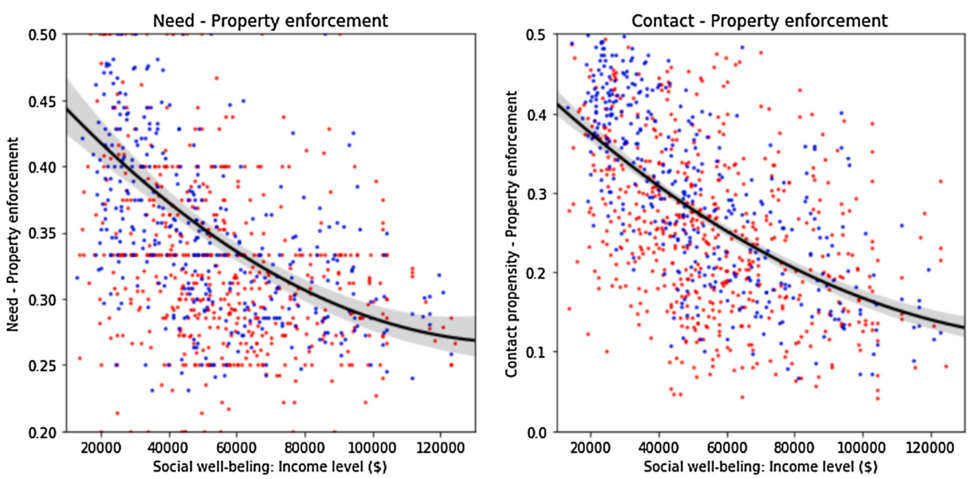

Fig. 6. Need and contacting propensity – code enforcement. (The need measure for property enforcement is the inverse value of resident satisfaction score of code enforcement. Colors are based on the classification results described in Fig. 4.)

Table 3
Pothole complaints and outcomes for under- and over-reporting neighborhoods in 2017. (Normalized values per street mile, per parcel are in parentheses.)

	Under-reporting	Over-reporting
No. of pothole complaints reported (mean)	2.51 (0.0056)	3.10 (0.058)
No. of unique pothole issues (mean)	2.42 (0.0054)	2.96 (0.047)
% of non-valid complaints reported (duplicates and out of city responsibilities)	3.58%	18.97%
Total no. of potholes fixed	644	985
No. of potholes fixed (mean)	2.40	2.91
No. of 311 complaints per fixed pothole	0.38	0.58
Median time to resolution (fixed)	4 days	4 days

neighborhoods by income quartile, as well as in the expected number of pothole repairs (Fig. 7). While the lowest income neighborhoods have the highest objective need (mean OCI of 49.76) and subjective need (satisfaction with city street infrastructure of 2.23 out of 5), there are far fewer repairs than would be expected. The OCI-based method corrects for this to a degree, but creates significant operational challenges given the time and cost of individual street visual inspections and its limitations as a lagging indicator of need.

Table 4
Expected pothole repairs by neighborhood income quartiles, based on three resource allocation methods: 311 complaints-only, street condition scores-only, and the proposed bias-adjusted 311 approach (2017 data).

	Q1	Q2	Q3	Q4
Neighborhood median income	$25,738	$42,977	$59,492	$97,885
Total expected potholes (311)	413.08	453.11	397.45	401.38
Total expected potholes (OCI)	543.20	425.43	377.73	318.76
Total expected potholes (bias-adjusted 311)	626.34	395.38	323.30	319.98
Expected potholes per mile (311)	1.06	1.57	1.27	1.32
Expected potholes per mile (OCI)	1.28	1.30	1.17	1.04
Expected potholes per mile (bias-adjusted 311)	1.35	1.07	0.90	0.95
Actual potholes fixed per mile (mean)	0.022	0.034	0.041	0.016
Resident satisfaction score (mean)	2.23	2.34	2.74	2.73
Street pavement score (OCI) (mean)	49.76	48.94	54.31	59.23

Our bias-adjusted 311 method provides a more equitable distribution of pothole repairs. In this case, the greatest number of potholes are repaired in the two lowest income quartiles. Since this approach

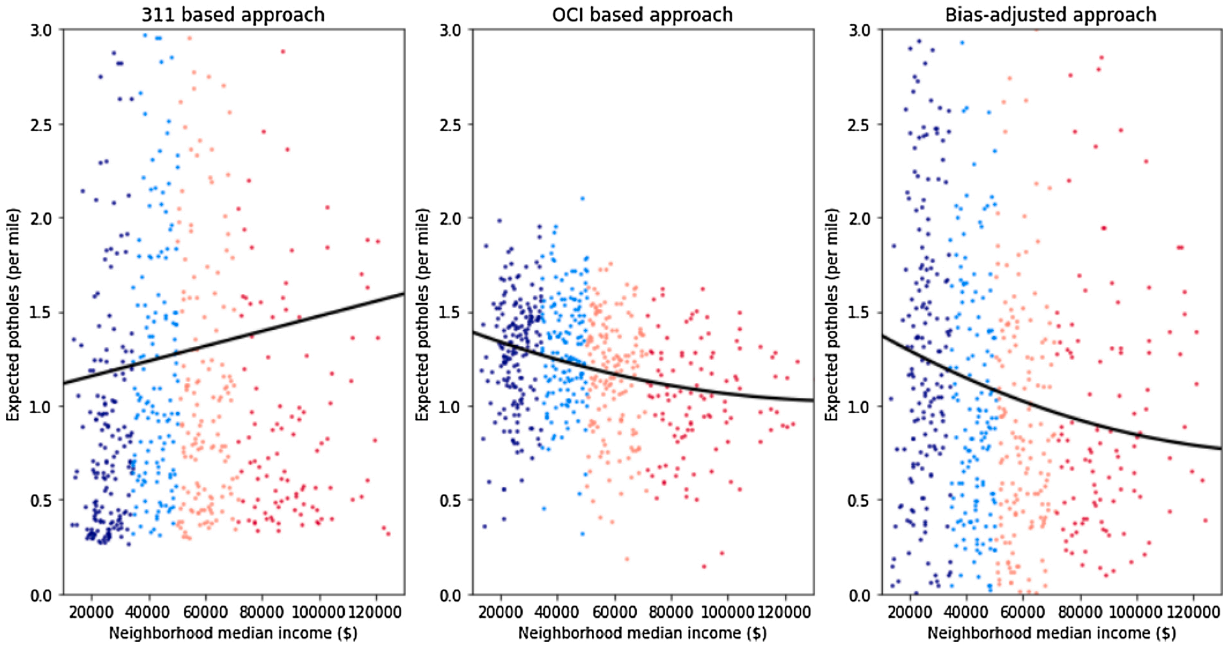

Fig. 7. Pothole repairs by neighborhood income based on three resource allocation methods: 311, street condition scores, and bias-adjusted 311. (Colors are neighborhood income quartiles.)

accounts for both reporting behavior and actual need, it represents a resource allocation method that can be used by local governments to assess dynamic (real-time) need through 311 reports, while recognizing that not all neighborhoods report these issues at the same rate when facing similar conditions.

6. Discussion and conclusion

We present an approach to identify disparities in 311 reporting based on household and neighborhood factors as evidence of potential bias in resident self-reported data. In previous work, explaining and quantifying the unequal use of 311 systems has been a challenge without objective validation data describing the location and extent of actual problematic conditions. This study overcomes these methodological challenges and data constraints by integrating 311 complaint reports with both actual and perceived measures of need related to street conditions. Resident-government interactions through the 311 system in Kansas City are found to vary with household demographic and socio-economic characteristics, levels of civic engagement and political participation, and reported satisfaction with local government infrastructure and services.

The case study demonstrates how our methodology can be used to identify under- and over-reporting neighborhoods based on actual street conditions across the city, creating an objective measure by which to estimate reporting propensities. Despite lower satisfaction levels and lower street maintenance quality, neighborhoods characterized by lower income, minority populations are less likely to report these “nuisance” issues. Given the presence of more severe life safety concerns in their communities, residents prioritize 311 reporting to address more pressing problems, such as hazards associated with property code enforcement issues. On the other hand, over-reporting neighborhoods complain more than expected about street conditions given their reported high level of satisfaction with local services and quality of local street infrastructure, suggesting that expectations for quality-of-life services are influenced by neighborhood context.

Based on this evidence, we demonstrate reporting bias in the nature and distribution of complaints received through 311. This bias stems from the representativeness of self-reported condition assessments, which are shaped by socio-cultural characteristics and neighborhood context. As such, using these data for data-driven city service delivery models can lead to an inequitable distribution of services resulting from the over-allocation of resources to households and neighborhoods that are more likely to report problems, despite observed discrepancies with actual conditions. We demonstrate this mis-allocation by showing where potholes would be repaired based on 311-only, condition assessment-only, and bias-adjusted 311 complaint resource allocation methods. The results indicate that our bias-adjusted method provides a more equitable distribution of repair services, while accounting for both reporting behavior differences across neighborhoods and actual conditions.

Although we focus on a specific complaint type in this study, our method can be applied to other problem types and is generalizable to other cities with appropriate data resources. Acquisition of more ground-truth data on other local problems will enable additional explorations into the spatial heterogeneity in resident-government interactions. As equity issues are discovered for other complaint types, city agencies can utilize our approach to identify and quantify bias in reported data and adjust resource allocation decision processes to account for both actual conditions and resident need. Since 311 data are shown not to be representative of the needs of the entire population, nor necessarily reflective of the nature and severity of actual problems and local conditions, city governments should be wary of the use of 311 data to predict, forecast, and respond to local problems. City agencies need to recognize the potential for, and magnitude of, sample bias in self-reported data resulting from differential reporting behavior. Our work forms the basis for acknowledging and accounting for this data bias, and can contribute to bias-aware public sector decision support systems. By using the proposed bias-adjustment method, city governments can improve resource allocation strategies for more efficient and equitable government services, taking into account neighborhood context and residents’ reporting propensities. In addition, the findings can be used by city agencies to encourage more widespread use of 311 through targeted outreach and community education for under-served residents and neighborhoods. Consequently, acknowledging and addressing bias in resident 311 complaints is a necessary starting point for more equitable, fair, and effective city service delivery through data-driven decision making processes.

Conflict of interest

The authors declare that there is no conflict of interest.

Funding

This material is based upon work supported by the National Science Foundation under award No. 1926470 and by a grant from the John D. and Catherine T. MacArthur Foundation. Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of any supporting institution. All errors remain the authors.

Declaration of Competing Interest

The authors report no declarations of competing interest.

Acknowledgements

The authors would like to thank DataKC for providing the data used in this study and insightful feedback on early versions of this work, as well as the five anonymous reviewers for their constructive comments. A preliminary version of this paper was presented at the 2018 Bloomberg Data for Good Exchange conference.

References

Amini, A., Soleimany, A. P., Schwarting, W., Bhatia, S. N., & Rus, D. (2019). Uncovering and mitigating algorithmic bias through learned latent structure. *Proceedings of the 2019 AAAI/ACM conference on AI, ethics, and society*, 289–295.

Barocas, S., Boyd, D., Friedler, S., & Wallach, H. (2017). *Social and technical trade-offs in data science*.

Batty, M. (2012). Smart cities, big data. *Environment and Planning B: Planning and Design, 39*, 191–193.

Bettencourt, L. M. (2014). The uses of big data in cities. *Big Data, 2*, 12–22.

Bouzguenda, I., Alalouch, C., & Fava, N. (2019). Towards smart sustainable cities: A review of the role digital citizen participation could play in advancing social sustainability. *Sustainable Cities and Society, 50*, 101627.

Brady, H. E., Verba, S., & Schlozman, K. L. (1995). Beyond SES: A resource model of political participation. *American Political Science Review, 89*, 271–294.

Cavallo, S., Lynch, J., & Scull, P. (2014). The digital divide in citizen-initiated government contacts: A GIS approach. *Journal of Urban Technology, 21*, 77–93.

Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data, 5*, 153–163.

Christensen, T., & Lægreid, P. (2005). Trust in government: The relative importance of service satisfaction, political factors, and demography. *Public Performance & Management Review, 28*, 487–511.

City of Kansas City, M. (2018). *Citizen and business satisfaction survey results*. Retrieved from http://kcmo.gov/data/citizen-satisfaction-survey-results/.

Drosou, M., Jagadish, H., Pitoura, E., & Stoyanovich, J. (2017). Diversity in big data: A review. *Big Data, 5*, 73–84.

Duvier, C., Anand, P. B., & Oltean-Dumbrava, C. (2018). Data quality and governance in a UK social housing initiative: Implications for smart sustainable cities. *Sustainable Cities and Society, 39*, 358–365.

Garcia, M. (2016). Racist in the machine: The disturbing implications of algorithmic bias. *World Policy Journal, 33*, 111–117.

Glaeser, E. L., Kominers, S. D., Luca, M., & Naik, N. (2018). Big data and big cities: The promises and limitations of improved measures of urban life. *Economic Inquiry, 56*, 114–137.

Hajian, S., Bonchi, F., & Castillo, C. (2016). Algorithmic bias: From discrimination discovery to fairness-aware data mining. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2125–2126.

Hartmann, S., Mainka, A., & Stock, W. G. (2017). Citizen relationship management in local governments: The potential of 311 for public service delivery. *Beyond bureaucracy* (pp. 337–353). Springer.

Hirlinger, M. W. (1992). Citizen-initiated contacting of local government officials: A multivariate explanation. *The Journal of Politics, 54*, 553–564.

Horgan, D., & Dimitrijević, B. (2019). Frameworks for citizens participation in planning: From conversational to smart tools. *Sustainable Cities and Society, 48*, 101550.

Huang, B., Xie, C., Tay, R., & Wu, B. (2009). Land-use-change modeling using unbalanced support-vector machines. *Environment and Planning B: Planning and Design, 36*, 398–416.

Ibrahim, S. A., Charlson, M. E., & Neill, D. B. (2020). Big data analytics and the struggle for equity in health care: The promise and perils. *Health Equity, 4*, 99–101.

Johnson, S. (2010). *What a hundred million calls to 311 reveal about New York*. https://www.wired.com/2010/11/ff_311_new_york/.

Jones, B. D., Greenberg, S. R., Kaufman, C., & Drew, J. (1977). Bureaucratic response to citizen-initiated contacts: Environmental enforcement in Detroit. *American Political Science Review, 71*, 148–165.

KCStat. (2015). *Infrastructure and transportation, March 10th 2015*. Retrieved from https://www.kcmo.gov/home/showdocument?id=190.

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal, 79*, 1–14.

Klosterman, R. E. (1994). Large-scale urban models retrospect and prospect. *Journal of the American Planning Association, 60*, 3–6.

Kontokosta, C. E. (2018). Urban informatics in the science and practice of planning. *Journal of Planning Education and Research*, 0739456X18793716.

Kontokosta, C., Hong, B., & Korsberg, K. (2017a). *Equity in 311 reporting: Understanding socio-spatial differentials in the propensity to complain*. arXiv preprint arXiv:1710.02452.

Kontokosta, C. E., Weiss, M., Snively, C., & Gulick, S. (2017b). *NYC311, Harvard Business School Case 818-056*.

Layne, K., & Lee, J. (2001). Developing fully functional e-government: A four stage model. *Government Information Quarterly, 18*, 122–136.

Levine, J. R., & Gershenson, C. (2014). From political to material inequality: Race, immigration, and requests for public goods. In *Sociological forum* (Vol. 29, pp. 607–627). Wiley Online Library.

McClure, D. L. (2000). *Statement of David L. Mcclure, US General Accounting Office, before the Subcommittee on Government Management, Information and Technology, Committee on Government Reform, House of Representatives*. Committee on Government Reform.

Melkers, J., & Thomas, J. C. (1998). What do administrators think citizens think? Administrator predictions as an adjunct to citizen surveys. *Public Administration Review*, 327–334.

Minkoff, S. L. (2016). NYC 311: A tract-level analysis of citizen-government contacting in New York City. *Urban Affairs Review, 52*, 211–246.

O'Brien, D. T. (2016a). 311 hotlines, territoriality, and the collaborative maintenance of the urban commons: Examining the intersection of a coproduction policy and evolved human behavior. *Evolutionary Behavioral Sciences, 10*, 123.

O'Brien, D. T. (2016b). Using small data to interpret big data: 311 reports as individual contributions to informal social control in urban neighborhoods. *Social Science Research, 59*, 83–96.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data, 1*, 51–59.

Schill, M. H., & Wachter, S. M. (1995). The spatial bias of federal housing law and policy: Concentrated poverty in urban America. *University of Pennsylvania Law Review, 143*, 1285–1342.

Schwester, R. W., Carrizales, T., & Holzer, M. (2009). An examination of the municipal 311 system. *International Journal of Organization Theory & Behavior, 12*, 218–236.

Sharp, E. B. (1984). Need, awareness, and contacting propensity: Study of a city with a central complaint unit. *Urban Affairs Quarterly, 20*, 22–30.

Sjoberg, F. M., Mellon, J., & Peixoto, T. (2017). The effect of bureaucratic responsiveness on citizen participation. *Public Administration Review, 77*, 340–351.

Teo, T. S., Srivastava, S. C., & Jiang, L. (2008). Trust and electronic government success: An empirical study. *Journal of Management Information Systems, 25*, 99–132.

Thijssen, P., & Van Dooren, W. (2016). Who you are/where you live: Do neighbourhood characteristics explain co-production? *International Review of Administrative Sciences, 82*, 88–109.

Thomas, J. C. (2013). Citizen, customer, partner: Rethinking the place of the public in public management. *Public Administration Review, 73*, 786–796.

Vedlitz, A., Dyer, J. A., & Durand, R. (1980). Citizen contacts with local governments: A comparative view. *American Journal of Political Science*, 50–67.

Verma, S., & Rubin, J. (2018). Fairness definitions explained. In *2018 IEEE/ACM international workshop on software fairness (FairWare)* (pp. 1–7).

Wang, C. Q. C. K., Lingjing, S., & Sobolevsky. (2016). *Structure of 311 service requests as a signature of urban location*. arXiv preprint arXiv:1611.06660.

Welch, E. W., Hinnant, C. C., & Moon, M. J. (2004). Linking citizen satisfaction with e-government and trust in government. *Journal of Public Administration Research and Theory, 15*, 371–391.

Westraadt, L., & Calitz, A. (2020). A modelling framework for integrated smart city planning and management. *Sustainable Cities and Society*, 102444.

White, A., & Trump, K.-S. (2018). The promises and pitfalls of 311 data. *Urban Affairs Review, 54*, 794–823.

Wu, W.-N. (2020). Determinants of citizen-generated data in a smart city: Analysis of open 311 user behavior. *Sustainable Cities and Society*, 102167.

Xu, L., Kwan, M.-P., McLafferty, S., & Wang, S. (2017). Predicting demand for 311 non-emergency municipal services: An adaptive space-time kernel approach. *Applied Geography, 89*, 133–141.

Yiftachel, O. (1989). Towards a new typology of urban planning theories. *Environment and Planning B: Planning and Design, 16*, 23–39.

Zha, Y., & Veloso, M. (2014). Profiling and prediction of non-emergency calls in New York City. In *Proceedings of the workshop on semantic cities: Beyond open data to models, standards and reasoning*.

Zhang, Z., & Neill, D. B. (2016). *Identifying significant predictive bias in classifiers*. arXiv preprint arXiv:1611.08292.