

# Unraveling Environmental Justice in Ambient PM<sub>2.5</sub> Exposure in Beijing: A Big Data Approach

Yanyan Xu,<sup>†,‡,¶</sup> Shan Jiang,<sup>§,||</sup> Ruiqi Li,<sup>⊥,#</sup> Jiang Zhang,<sup>⊥</sup> Jinhua Zhao,<sup>@</sup>

Sofiane Abbar,<sup>△</sup> and Marta C. González<sup>\*,†,‡,¶</sup>

<sup>†</sup>*Department of Civil & Environmental Engineering, MIT, Cambridge, MA 02139, USA*

<sup>‡</sup>*Department of City and Regional Planning, University of California, Berkeley, CA 94720, USA*

<sup>¶</sup>*Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

<sup>§</sup>*Department of Urban and Environmental Policy and Planning, Tufts University, Medford, MA 02155, USA*

<sup>||</sup>*Department of Civil and Environmental Engineering, Tufts University, Medford, MA 02155, USA*

<sup>⊥</sup>*School of Systems Science, Beijing Normal University, Beijing 100875, China*

<sup>#</sup>*College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China*

<sup>@</sup>*Department of Urban Studies & Planning, MIT, Cambridge, MA 02139, USA*

<sup>△</sup>*Qatar Computing Research Institute, HBKU, Doha 5825, Qatar*

E-mail: martag@berkeley.edu

## Abstract

Air pollution imposes major environmental and health risks worldwide and is expected to worsen in the coming decade as cities expand. Measuring population exposure to air pollution is crucial to quantify the risks on health. In this work, we introduce

5 a data analysis framework to model residents' stay and commuters' travel exposure to  
6 outdoor PM<sub>2.5</sub> in Beijing and analyze their environmental justice in relation to wealth,  
7 for which the housing price is harnessed as a proxy. Using mobile phone and census  
8 data, we first infer the travel demand of population to achieve the dynamic stay in each  
9 zone, we then separate the commuters from all travelers and estimate their travel routes  
10 with a traffic assignment model. With the observations from air quality monitoring  
11 stations, we estimate the outdoor PM<sub>2.5</sub> concentration at a 500-meter grid level with  
12 a spatial interpolation model, and map the concentration to the road networks with  
13 the consideration of the average travel time on each road segment. By combining the  
14 estimated exposure and housing price, we discover that in the winter, commuters with  
15 low wealth level are exposed to 13% more PM<sub>2.5</sub> per hour than those with high wealth  
16 level when staying at home, but exposed to less PM<sub>2.5</sub> by 5% when commuting the  
17 same travel distance due to lighter traffic congestion in suburban areas. We also unveil  
18 that the residents from the southern suburbs of Beijing have both lower wealth and  
19 higher stay- and travel- exposure to PM<sub>2.5</sub> especially in the winter due to the heavy  
20 coal burning in the south of Beijing. This finding informs policymakers to develop  
21 more equitable environmental mitigation policies for future sustainable development  
22 in Beijing. Additionally, for the first time in the literature, we compare the results of  
23 exposure estimated from passive data with subjective measures of perceived air quality  
24 (PAQ) obtained from a mobile-app based survey and shows both consistency in the  
25 results and the advantages of massive data over surveys, for air pollution exposure  
26 assessments.

## 27 **Keywords**

28 Environmental Justice, PM<sub>2.5</sub> Exposure, Travel Exposure, Urban Mobility, Mobile Phone  
29 Data

<sup>30</sup> **1. Introduction**

<sup>31</sup> With the rapid urbanization and the acceleration of industrialization, today's air pollution  
<sup>32</sup> has become a global threat to human health, especially for the large-scale and densely popu-  
<sup>33</sup> lated cities in developing countries.<sup>1–5</sup> According to the World Health Organization (WHO),  
<sup>34</sup> in 2012 around 3 million people died as a result of ambient air pollution exposure, which  
<sup>35</sup> makes it the largest environmental risk to the health of human beings worldwide. The partic-  
<sup>36</sup> ulate matter of a diameter of less than  $2.5 \mu\text{m}$  ( $\text{PM}_{2.5}$ ) is the major concern to the public in  
<sup>37</sup> recent years, especially for cities suffering from severe hazes. Taking Beijing as an example,  
<sup>38</sup> the outdoor  $\text{PM}_{2.5}$  concentration of 179 days of the year in 2015 was higher than  $75 \mu\text{g}/\text{m}^3$   
<sup>39</sup> and the average annual concentration reached  $80.6 \mu\text{g}/\text{m}^3$ .<sup>6</sup> Such concentrations represent  
<sup>40</sup> serious threats to human health, particularly to those vulnerable to heart or respiratory  
<sup>41</sup> diseases, such as the young and elderly population.<sup>7</sup>

<sup>42</sup> Due to these threats to health, quantifying human exposure to air pollutants have re-  
<sup>43</sup> ceived considerable attention and various types of data and methods have been introduced to  
<sup>44</sup> estimate their exposure in space and time.<sup>8–15</sup> Many previous studies infer pollutants concen-  
<sup>45</sup> tration from stationary air quality monitoring networks using spatial interpolation techniques  
<sup>46</sup> and estimate air pollution exposure at aggregated level with the spatial distribution of the  
<sup>47</sup> population.<sup>9</sup> This approach ignores human mobility and the time spent at various places in a  
<sup>48</sup> day.<sup>16,17</sup> The development of GPS-enabled mobile monitors and location-aware instruments  
<sup>49</sup> addresses this issue by measuring individual exposure during their activities.<sup>10,18–20</sup> Travel  
<sup>50</sup> survey data were also employed to examine the dynamic exposure of participating interview-  
<sup>51</sup> ees.<sup>12</sup> However, such datasets can only be used to estimate exposure of limited groups of  
<sup>52</sup> individuals, and it is hard to extrapolate the samples to the entire population.

<sup>53</sup> More recently, researchers have employed new data sources and methods to study the  
<sup>54</sup> dynamic exposure to air pollutants at the city scale.<sup>13,14,21</sup> For example, based on travel sur-  
<sup>55</sup>vey data, Shekarrizfard et al. developed an integrated transportation and emission model to  
<sup>56</sup> assess the individual stay and travel exposure to  $\text{NO}_2$  in Montreal, Canada.<sup>22</sup> Saraswat et al.

57 simulated the travel demand of population in New Delhi with a Gravity model and estimated  
58 the human exposure for different activities (such as staying at home, work and commuting  
59 in the road network).<sup>21</sup> However, they neglected the routing behavior and speed profiles of  
60 travelers when estimating the travel exposures. With the development of information and  
61 communication technologies (ICT), large scale geo-located mobile phone data have been in-  
62 creasingly used to model human mobility in cities.<sup>23,24</sup> For instance, Nyhan and colleagues  
63 have adopted mobile phone data to model the active population weighted exposure to PM<sub>2.5</sub>  
64 at aggregated levels in New York City,<sup>13</sup> and quantified the individual exposure to PM<sub>2.5</sub> by  
65 considering both home and work location.<sup>25</sup> Dewulf et al. used mobile phone data to track  
66 a user's visited locations and time of stays to estimate the dynamic exposure at individual  
67 level in Belgium.<sup>14</sup> Although these works took advantages of the large scale that mobile  
68 phone data could offer, they only considered the samples without extrapolating them to the  
69 population at the city level.<sup>26,27</sup> By only focusing on the exposures of stationary activities,  
70 these studies ignored human travels in cities. With the increased traffic congestion and long  
71 distance travels in large cities, people have spent more time on the road network than ever  
72 before.<sup>28</sup> In consequence, people are exposed to non-negligible air pollution while traveling.  
73 Therefore, it is important to estimate the air pollution exposure on the road networks. More  
74 importantly, by estimating exposure to PM<sub>2.5</sub>, it will offer opportunities for researchers to  
75 examine the environmental justice for the economically disadvantaged population.<sup>29</sup> The  
76 evaluation of environmental justice in cities will be useful to inform policymakers to develop  
77 equitable strategies for sustainable urban futures.<sup>30</sup>

78 Towards this end, in this study by using passively collected big data, we present a frame-  
79 work to evaluate the environmental justice of PM<sub>2.5</sub> exposure for commuters of different  
80 wealth levels in Beijing from the urban mobility perspective. First, we infer the seasonally  
81 average PM<sub>2.5</sub> concentration per hour at a 500-meter grid level with air monitor observa-  
82 tions and a spatial interpolation algorithm. We then map the PM<sub>2.5</sub> concentration in the  
83 road networks. Second, we use mobile phone data and census data to efficiently model the

<sup>84</sup> citywide human mobility at both individual and aggregated levels without the expensive  
<sup>85</sup> travel survey.<sup>23,31-34</sup> With a traffic assignment model, we estimate individual stay location,  
<sup>86</sup> duration, travel routes and travel time. Subsequently, we estimate the residents' outdoor  
<sup>87</sup> stay exposure to PM<sub>2.5</sub> by weighting population density. We also model the travel exposure  
<sup>88</sup> of commuters, by accumulating the PM<sub>2.5</sub> exposure on each traversed road segment, taking  
<sup>89</sup> into consideration travel time in traffic. Third, we investigate the environmental justice of  
<sup>90</sup> residents by connecting human exposure to PM<sub>2.5</sub> with housing prices (a proxy for wealth) in  
<sup>91</sup> Beijing, which is important for policymakers to develop equitable environmental mitigation  
<sup>92</sup> policies for the city. Finally, by comparing our results with a mobile-phone based survey  
<sup>93</sup> on individual perception of air quality, we assess the feasibility of using large-scale mobile  
<sup>94</sup> phone data to measure human exposure to the air pollution in the city.

## <sup>95</sup> 2. Materials and methods

### <sup>96</sup> 2.1 Estimating the spatial concentration of PM<sub>2.5</sub>

<sup>97</sup> As one of the most crowded cities in the world, Beijing accommodates 21.5 million residents  
<sup>98</sup> in an area of 16,410 km<sup>2</sup>. Since 75% of Beijing residents live in the urban area within the  
<sup>99</sup> Sixth Ring Road, a 20% of the total land area of the Beijing metropolitan area (Figure 1A),  
<sup>100</sup> this study is focused in this area. Beijing Municipal Environmental Monitoring Center  
<sup>101</sup> (BJMEMC)<sup>6</sup> collects concentrations of major air pollutants on an hourly basis from 35  
<sup>102</sup> air quality monitoring stations, among which 24 are located within the Sixth Ring Road  
<sup>103</sup> (Fig. 1A). The PM<sub>2.5</sub> concentration in Beijing displays strong seasonality, in both climate  
<sup>104</sup> characteristics and economic activities such as coal heating in the winter.<sup>35,36</sup> To derive  
<sup>105</sup> the representative seasonal air quality data, we average the PM<sub>2.5</sub> concentration for each  
<sup>106</sup> monitoring station during a given hour in the summer (from June 1st to August 31st, 2015)  
<sup>107</sup> and winter (from December 1st, 2015 to February 28th, 2016), respectively. Fig. 1B shows  
<sup>108</sup> the average PM<sub>2.5</sub> concentration per hour for the 24 stations, and their respective average

<sup>109</sup> value in the summer and winter. The average PM<sub>2.5</sub> concentrations in the summer were  
<sup>110</sup> stable and below 75  $\mu\text{g}/\text{m}^3$ ; while in the winter, the values were relatively higher at night  
<sup>111</sup> than those during the day. The PM<sub>2.5</sub> concentrations in the winter at most monitoring  
<sup>112</sup> stations were higher than 100  $\mu\text{g}/\text{m}^3$  and ranked as either unhealthy or very unhealthy  
<sup>113</sup> based on standards defined by the U.S. Environmental Protection Agency (EPA).<sup>37</sup>

<sup>114</sup> Estimating the PM<sub>2.5</sub> spatial concentrations is the first step to quantify human exposure  
<sup>115</sup> to air pollutants in the city. By applying the Ordinary Kriging (OK) spatial interpolation  
<sup>116</sup> method,<sup>38,39</sup> we estimate the PM<sub>2.5</sub> concentration at a 500 meter grid level in sub-regions of  
<sup>117</sup> the city based on the seasonal average values of the 24 monitoring stations per hour. Another  
<sup>118</sup> widely used method for mapping the spatial PM<sub>2.5</sub> concentration with sparse monitoring data  
<sup>119</sup> is land use regression (LUR) model, which utilizes the geographic characteristics to refine  
<sup>120</sup> the estimation of PM<sub>2.5</sub> concentration. Zou et al. evaluated the performance of LUR model  
<sup>121</sup> and OK interpolation in Houston and revealed that LUR and OK have similar end-result  
<sup>122</sup> accuracy for their comparable error rates, 6.13% and 7.01%, respectively, in the estimation  
<sup>123</sup> of PM<sub>2.5</sub> concentration.<sup>39</sup>

<sup>124</sup> OK predicts the values at the target locations based on the distance and spatial distri-  
<sup>125</sup> bution of the target location. Each of the (500 meter by 500 meter) grid is then assigned  
<sup>126</sup> a PM<sub>2.5</sub> concentration per hour for the summer and winter. Fig. 1C exhibits the estimated  
<sup>127</sup> PM<sub>2.5</sub> concentration in each zone *Jiedao* (a community-level zone in China) during the morn-  
<sup>128</sup> ing peak hour, the midday off-peak hour, and the evening peak hour in both summer and  
<sup>129</sup> winter. As shown in the figure, PM<sub>2.5</sub> in the winter were more concentrated in the southeast  
<sup>130</sup> of Beijing, where industrial activities played a dominant role.

## <sup>131</sup> 2.2 Inferring urban mobility from mobile phone data

<sup>132</sup> The travel demand for the 16.3 million residents living in the urban area of Beijing is esti-  
<sup>133</sup> mated using mobile phone data. The mobile phone dataset contains 100,000 active users  
<sup>134</sup> with their call detail records (CDRs) and data detail records (DDRs) for December 2013.

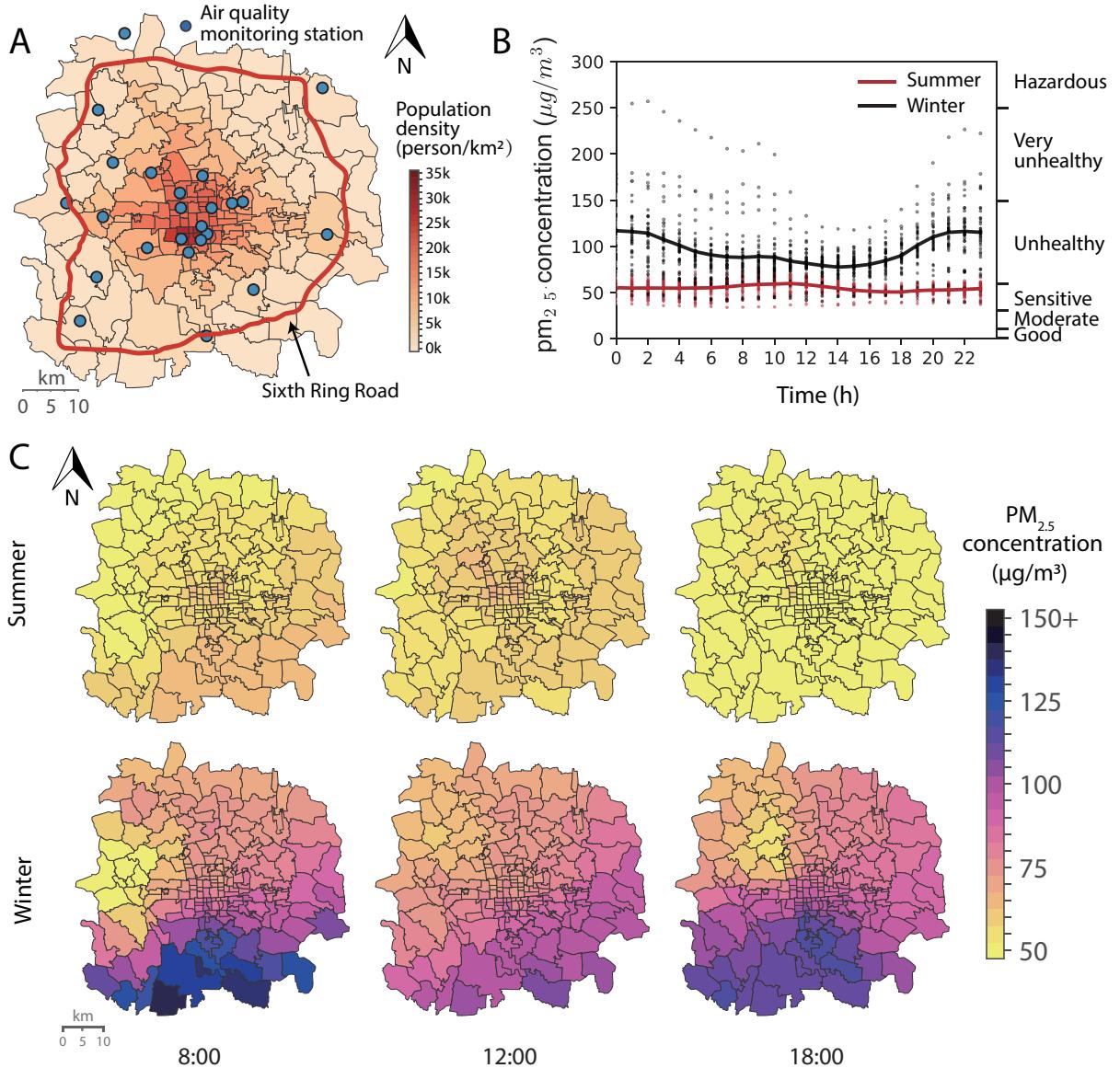


Figure 1: **Observed and estimated PM<sub>2.5</sub> concentration in Beijing.** **A** Population density and air quality monitoring stations within the Sixth Ring Road of Beijing. **B** Average hourly PM<sub>2.5</sub> concentration during the summer and winter of 2015. The scatter dots denote the concentration of the monitoring stations and the solid line shows the average concentration of all monitoring stations. **C** Estimated average PM<sub>2.5</sub> concentration in each zone at selected hours, 8:00, 12:00 and 18:00, in summer and winter. [Figure created with Basemap Matplotlib Toolkit for Python]

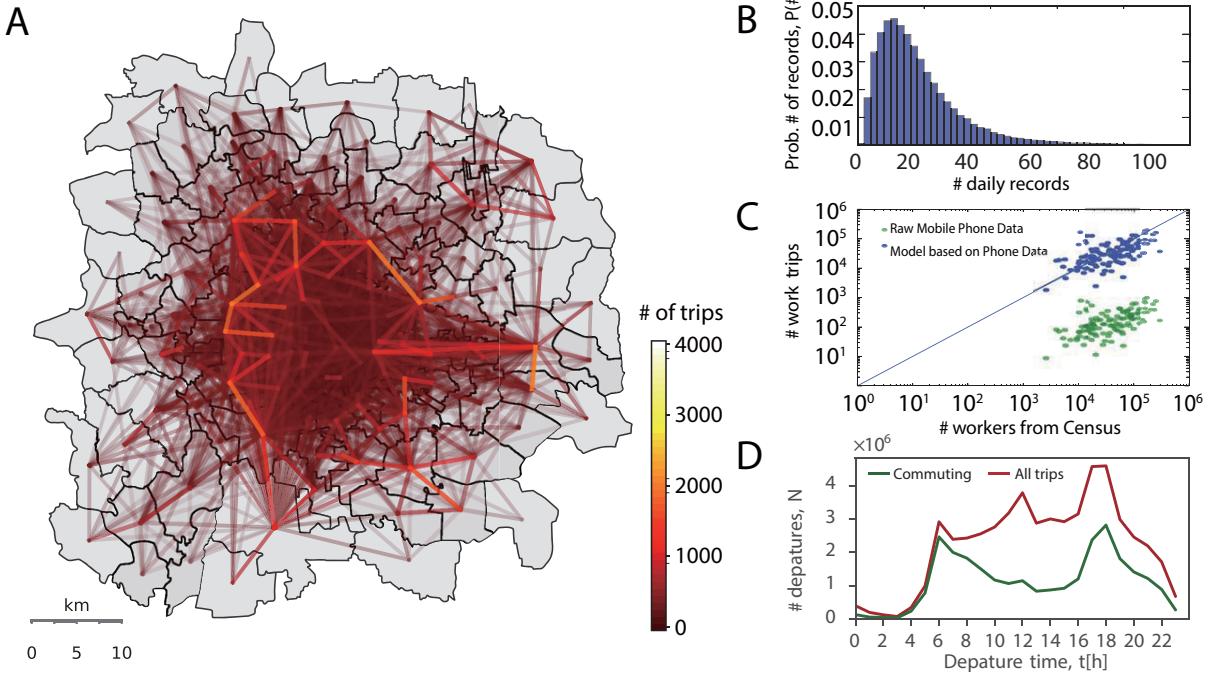
135 The communication activities of CDRs comprise the incoming and outgoing phone call and  
 136 the sending and receiving of a text message, while DDRs comprise the using of Internet  
 137 data. Each record of the CDRs and DDRs data has a hashed ID, start time-stamp of the

138 activity, type of activity, duration of the activity, longitude and latitude of the cell tower  
139 that communicated with the phone. The hashed ID is unique for each mobile phone user,  
140 so that we are able to analyze the anonymized user when she is interacting with the phone  
141 which communicates with the nearest cell tower. In rare cases, the second nearest tower  
142 will be used if the nearest one is fully loaded. The cell phone in use will be switched to  
143 the closest cell tower when the user is moving. The average distance between cell towers is  
144 332 meters (with a median of 254 meters), representing the spatial resolution of the study.  
145 Alexander *et al.* and Colak *et al.* outlined a general framework to obtain the travel demand,  
146 a.k.a. Origin-Destination (OD) matrices, from massive mobile phone data.<sup>40,41</sup> We apply  
147 the same method here to extract daily trips from mobile phone samples in Beijing. These  
148 trips are then combined with the resident population of each zone within the Sixth Ring  
149 Road to estimate representative urban travel demand at the zonal level. Before modeling  
150 the travel behavior of each anonymous users, we first eliminate non active users. Then,  
151 we extract stay locations of the active users from their raw records. We improve upon the  
152 stay-point algorithm presented by Zheng and Xie<sup>42</sup> and Jiang et al.<sup>43</sup> as follows: (1) We  
153 apply a temporal agglomeration algorithm, such that temporally consecutive records within  
154 a certain radius (e.g., 500 meters) are bundled together with an updated stay duration from  
155 the start time of the first record to the end time of last one; (2) We label the records as pass-  
156 by points and stays, based on the stay duration threshold (e.g., 10 minutes). In the analysis  
157 hereafter, we only focus on the stays. We then combine all the spatially adjacent stay points  
158 for a user (within a threshold) as his or her stay regions. For this spatial agglomeration,  
159 we use a spatial search balancing tree, R-tree, to accelerate the computation (see Fig. S1 in  
160 Supplementary Material (SM)<sup>44</sup>).

161 After the stay locations are detected for each user, we label their stays as *home*, *work*, or  
162 *other*. The most frequently visited location during weekday nights and weekends are labeled  
163 as *home*, and the most frequently visited one during weekday working hours (at least 500  
164 meters away from home) is labeled as *work*, if one exists, and the rest are labeled as *other*.

165 Each trip can be then labeled as one of the three categories, (1) home-based-work (HBW),  
166 which refers to the trips between home and work, a.k.a. commuting; (2) home-based-other  
167 (HBO), which refers to the trips between home and other places; (3) non-home-based (NHB),  
168 which refers to the trips between work and other places or two other places. Among the  
169 three categories of trips, commuting flow among zones are the most stable as the urban  
170 population and economic structure are relatively stable in the metropolitan area, such as  
171 Beijing. Eventually, as the mobile phone users only cover a part of the entire population, we  
172 expand the travel portrait of the mobile phone samples with expansion factor by zone. The  
173 expansion factor is defined as the ratio between the actual resident population from census  
174 data and the number of mobile phone users whose *home* are located in the given zone. After  
175 aggregating the trips at the zone level, we estimate the OD matrices by hour for an average  
176 weekday and 24 OD matrices were derived for the population and commuters, respectively.

177 Fig. 2A shows the OD pairs with the large commuting flows between zones for the morning  
178 peak hour, obtained from the above discussed mobility model. Fig. 2B shows the average  
179 number of phone usage records per day per user during a month. Majority of users are active,  
180 with an average of 15 daily records. As we observe the travel demand of individuals during  
181 one month, the records cover most places that they visited in their daily life, especially  
182 the work places for commuters. To validate the estimated commuting travel demand, we  
183 compare with the census employment statistics at the *Jiedao* level,<sup>45</sup> and find that our  
184 employment estimation is in good agreement with the Beijing 2<sup>nd</sup> Economic Census (see  
185 Fig. 2C). The number of commuting and all trips per hour are shown in Fig. 2D. The  
186 morning and evening peaks can be observed from the commuting trips, while there are three  
187 peaks for all trips during the morning, mid-day and evening on an average weekday. While  
188 a travel survey from Beijing is not available for this study, this method has been validated  
189 in many other cities with travel surveys, and traditional travel demand models developed by  
190 planning agencies.<sup>40,41,46</sup> More details of the urban mobility model results can be found in  
191 the Note in SM.



**Figure 2: Estimation and validation of urban mobility in Beijing.** **A** The top 5,000 origin-destination commuting flows among zones during the weekday morning peak hour. **B** Distribution of the number of daily records of each active user in the mobile phone dataset. Most users interact with the mobile phone for 10–30 times per day, including phone calls, text messages and Internet data. **C** Number of workers in each zone before and after expansion from users to the total population. The green circles show the comparison between recognized work trips from mobile phone data and the number of workers from census data. The blue circles show the number of work trips after expanding the mobile phone data to the whole population. **D** Number of trip departures of commuters and all population by time of the day.

As previous studies have shown that exposure to PM<sub>2.5</sub> can also be positively associated with increased psychological distress and affect human health,<sup>47</sup> to compare the objective estimate with individual subjective perception of air quality, we collected a smart-phone based survey from individuals in Beijing in this study. To our knowledge, this is the first survey in a Chinese city that is dedicated to capturing the perceived air quality by local residents. In the perceived air quality (PAQ) experiment, more than 26,000 individuals received the study invitation, and around 1,000 individuals expressed interest to participate. Among those, 860 individuals downloaded our smartphone-based survey application to track their daily trajectory, and 256 of them finished the survey by rating the PAQ for

201 home, workplace, and worst spot during their commuting during the two week study period  
202 in the winter of 2015 (although some of them didn't complete the whole period). The  
203 perception runs from 0 to 5, with 0 being the best air quality and 5 being the worst. When  
204 comparing the daily PM<sub>2.5</sub> with the daily average PAQ for these respondents during the  
205 study period, we find that they are highly correlated with a Pearson correlation coefficient  
206 of 0.831.

## 207 2.3 Modeling stay and travel exposure to PM<sub>2.5</sub>

208 With our inferred mobility at the urban population scale, we account the hourly dynamic  
209 stay of the population and the travel time and route choice of commuters assuming all  
210 residents starts their daily trips from home. We then estimate the stay exposure to PM<sub>2.5</sub>  
211 of population and the travel exposure to PM<sub>2.5</sub> of commuters per hour in the summer and  
212 winter, respectively. We define population density weighted exposure (PDWE) to represent  
213 the total outdoor exposure of population per square kilometer per hour. For a given zone  
214  $z$  during an hour  $h$ , its PDWE is defined as  $E_{h,z} = C_{h,z} \cdot P_{h,z}/S_z$ , where  $C_{h,z}$  denotes the  
215 PM<sub>2.5</sub> concentration during that hour,  $P_{h,z}$  denotes the dynamic population staying in the  
216 zone during the same hour and  $S_z$  denotes the area of the zone. The unit of PDWE is  
217  $\text{person} \cdot \mu\text{g} \cdot \text{m}^{-3} \cdot \text{km}^{-2} \cdot \text{h}$ . Unlike the population weighted exposure (PWE) introduced by  
218 Nyhan et al.,<sup>13</sup> which is density-independent and tied to the total population of the zone,  
219 PDWE highlights the areas with higher population density and heavier air pollution. As  
220 shown in Fig. 1A, the size of zones are widely different, which causes the total population of  
221 low density zones in suburban area is higher than that of high density zones in the central  
222 area. However, air pollution impose more serious threats to zones with high density.

223 The travel exposure to PM<sub>2.5</sub> of a commuter depends on her selected route, travel time,  
224 and PM<sub>2.5</sub> concentration on the road segments. In this work, we focus on commuters who  
225 travel between zones either with cars or buses as they are the main travel modes for long-  
226 distance above-ground commuting in Beijing. Based on the travel demand of commuters

227 inferred from mobile phone data and census data, we use the travel mode share reported in  
228 the Travel Survey of Beijing Residents to estimate trips made by passenger cars and buses  
229 per hour.<sup>48,49</sup> Regarding the traffic conditions in the road networks, we derived vehicle trips  
230 from the vehicle usage rate per origin zone. We then estimated the driving time per route  
231 by assigning the vehicle OD matrices to the road network using a traffic assignment model  
232 based on user equilibrium (UE). Further details are discussed in Ref.<sup>41,50</sup> The road network  
233 is extracted from OpenStreetMap.<sup>51</sup> The UE model gives each OD pair its shortest travel  
234 time and paths. We validated our estimates of travel times with Gaode Map,<sup>52</sup> a widely  
235 used travel navigation platform in China. The comparison of the distribution of travel  
236 time indicates good estimates, shown in Fig. S2B and .S2C in SM. The Pearson correlation  
237 coefficient between the travel times estimated by our model and Gaode Map is 0.84 and the  
238 relative accuracy is 79.17% by regarding Gaode Map as the ground truth. For simplicity,  
239 we assume that buses use similar routes as cars, but with longer travel times. According to  
240 a travel survey report, the travel time for a trip made by bus equals to 1.57 times for a car  
241 trip on average.<sup>48</sup>

242 With the traffic assignment model, each road segment in the road network is associated  
243 with a travel time with traffic. Further, we can estimate the PM<sub>2.5</sub> concentration on each  
244 road segment by mapping the grid map of PM<sub>2.5</sub> concentration to the roads. When a road  
245 segment covers more than one grid, the average concentration of all covered grids is regarded  
246 as the PM<sub>2.5</sub> concentration of the road segment (see Fig. S3 in the SM). The estimated PM<sub>2.5</sub>  
247 concentration of the road network during the morning peak hour, mid-day, and the evening  
248 peak hour in summer and winter are shown in Fig. S4 in the SM. Finally, we calculate the  
249 route travel exposure to PM<sub>2.5</sub> of a commuter by aggregating the road segment exposure,  
250  $E_h^T = \sum_{r=1}^R C_h^r T_h^r$ , where  $C_h^r$  and  $T_h^r$  denote respectively the PM<sub>2.5</sub> concentration and travel  
251 time on the  $r$ th road segment in the route during the  $h$ th hour and  $R$  is the number of road  
252 segments forming the route. The unit of travel exposure is  $\mu\text{g} \cdot \text{m}^{-3} \cdot \text{h}$ .

253 **3. Results**

254 **3.1 Population density weighted exposure**

255 As most travel activities happen during the daytime (*e.g.*, work, school), we calculate the  
 256 stay exposure to PM<sub>2.5</sub> in two periods, work hours and non-work hours. The former covers  
 257 hours between 9:00 am to 5:00 pm; the latter covers the rest of the day. Fig. 3A illustrates  
 258 the average hourly PDWE during non-work and work hours in the summer and winter. By  
 259 observing the spatial distribution of PDWE, we identify exposure by zones during non-work  
 260 and work hours. The disparity of PDWE between work and non-work hour is mainly caused  
 261 by the travel activities of residents, and the disparity between summer and winter is mainly  
 262 caused by the seasonal variations in PM<sub>2.5</sub> concentration in Beijing. Specifically, some central

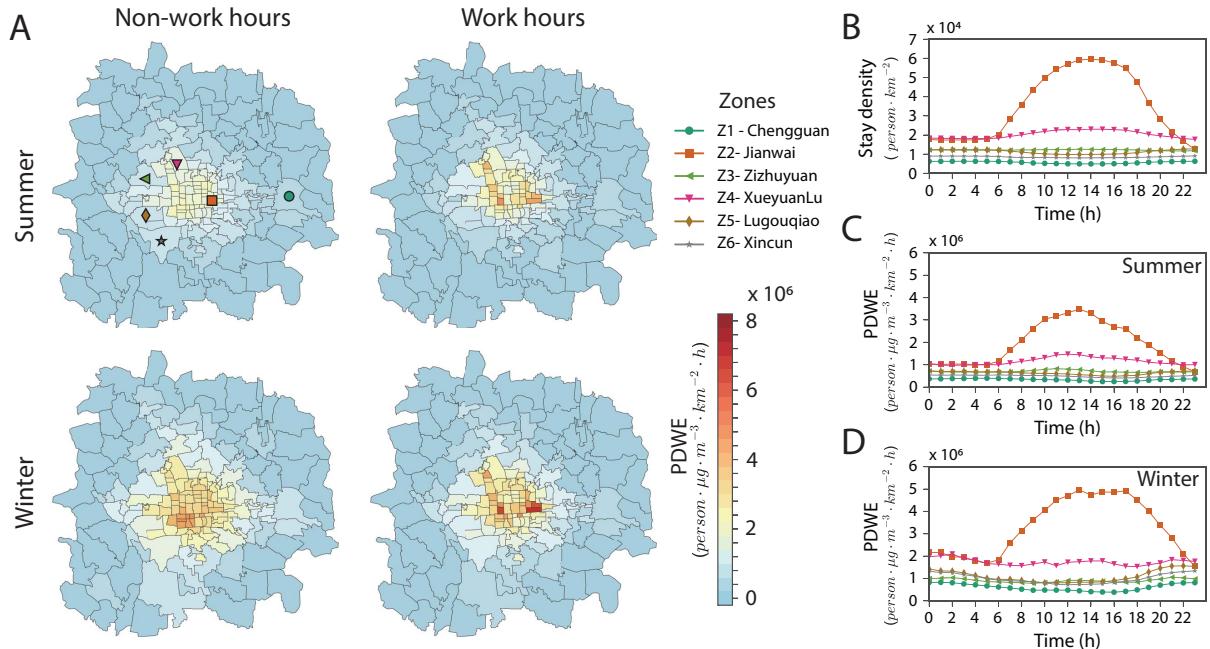


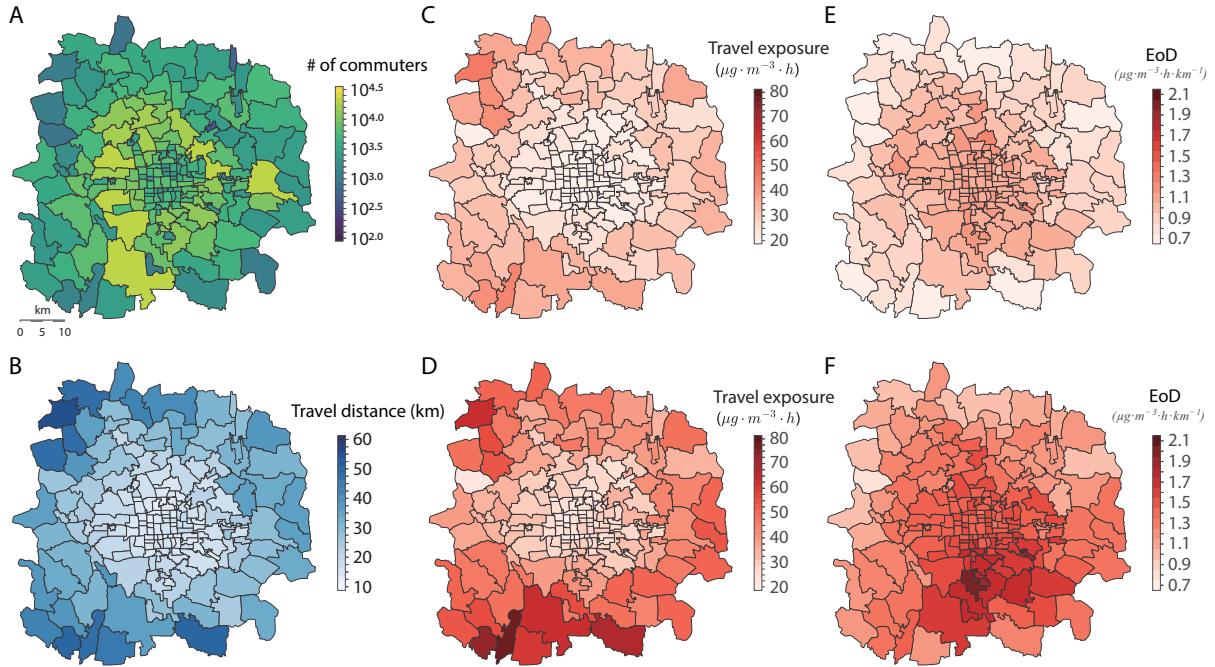
Figure 3: **Population density weighted exposure (PDWE) to PM<sub>2.5</sub> per zone.** **A** Average PDWE in each zone during work (from 9:00 am to 5:00 pm) and non-work hours during the summer and winter of 2015. The zones in downtown are exposed to heavier PM<sub>2.5</sub> concentration during work hours than non-work hours. **B** Hourly population density of six selected zones. The higher population density during daytime in Zone Z2 and Z4 captures the incoming flow at work hours. **C** Hourly PDWE of selected zones in the summer. **D** Hourly PDWE of selected zones in the winter.

263 areas in Beijing experience more severe PDWE during work hours than non-work hours as  
264 a significant portion of the population is gathering into the central area of the city during  
265 daytime for work and/or business. The disparity of stay exposure in space and time is more  
266 evident in winter than in summer.

267 According to the urban mobility patterns, we select six representative zones in the city,  
268 marked with different symbols in Fig. 3A, to uncover the PDWE. Fig. 3B displays the  
269 population density of stays per hour in the selected zones (keeping the colors of the labeled  
270 zones). The population density at noon is three times that of midnight in zone Z2, which is  
271 located in the CBD of Beijing, the Chaoyang district. Figs. 3C and 3D depicts the hourly  
272 PDWE of six selected zones during summer and winter, respectively. In the summer, the  
273 worst PDWE in zone Z2 reaches  $3.5 \times 10^6 \text{ person} \cdot \mu\text{g} \cdot m^{-3} \cdot km^{-2} \cdot h$ . While in winter it  
274 reaches  $5.0 \times 10^6 \text{ person} \cdot \mu\text{g} \cdot m^{-3} \cdot km^{-2} \cdot h$ .

### 275 3.2 Spatial variation of travel exposure to PM<sub>2.5</sub>

276 The travel exposure captures the air pollution for each commuter between home and work  
277 by car or bus. By combining the estimated traffic flow, route, and travel times between  
278 each OD pairs with the PM<sub>2.5</sub> concentration of the road network, we can estimate the travel  
279 exposure to PM<sub>2.5</sub> between any two zones in any hour of the day. We select the trips  
280 between 8:00 am and 9:00 am, which reflects commuters' trips and the spatial differences  
281 in their travel exposure. The map in Fig. 4A depicts the number of commuters that travel  
282 across zones during the morning peak hour. In Fig. 4B we show the average travel distance  
283 of commuters living in each zone. As expected, suburban areas display longer commuting  
284 distance as most jobs are centralized in the city center. Fig. 4C and 4D display the average  
285 travel exposure of commuters in each zone during the morning peak hour in the summer  
286 and winter, respectively. The commuting exposure in the summer and winter show some  
287 discrepancies, especially in the southern area where commuters experience higher travel  
288 exposure to PM<sub>2.5</sub> in the winter.



**Figure 4: Travel exposure to  $\text{PM}_{2.5}$  of commuters from each zone during the morning peak hour in summer and winter.** **A** Number commutes departures from each zone during the morning peak hour. **B** Average travel distance of commuters in each zone during the morning peak hour. **C, D** Average travel exposure to  $\text{PM}_{2.5}$  of commuters during the morning peak hour in summer and winter, respectively. **E, F** Exposure-over-distance of commuters during the morning peak hour in summer and winter, respectively.

289 To better evaluate the spatial variation of travel exposure to  $\text{PM}_{2.5}$ , we define the travel

290 exposure per kilometer in each zone as the fraction between the total travel exposure and

291 the total travel distance for commuters in a given zone, namely exposure-over-distance ratio

292 ( $\text{EoD}$ ,  $\mu\text{g} \cdot \text{m}^{-3} \cdot \text{h} \cdot \text{km}^{-1}$ ).  $\text{EoD}$  indicates the concentration of  $\text{PM}_{2.5}$  exposures to the

293 traveler per kilometer from the origin to the destination. Trips with larger  $\text{EoD}$  are exposed

294 to more  $\text{PM}_{2.5}$  than others even when they have the same commuting distance. Fig. 4E

295 and 4F illustrate the  $\text{EoD}$  per zone in summer and winter, respectively. In the summer,  $\text{EoD}$

296 displays higher values near the central area and lower values in the suburbs. This is caused

297 by the heavy traffic congestion in the central area, as shown in Fig. 2A in the SM. As for

298 the distribution of  $\text{EoD}$  in winter, the regions with highest values moves to the south due to

299 the combined effect of both higher  $\text{PM}_{2.5}$  concentration in the south (as shown in Fig. 1C)

300 and the heavy traffic congestion in the central area. More results are presented in Fig. S5 in

301 the SM and an on-line travel exposure visualization platform\*.

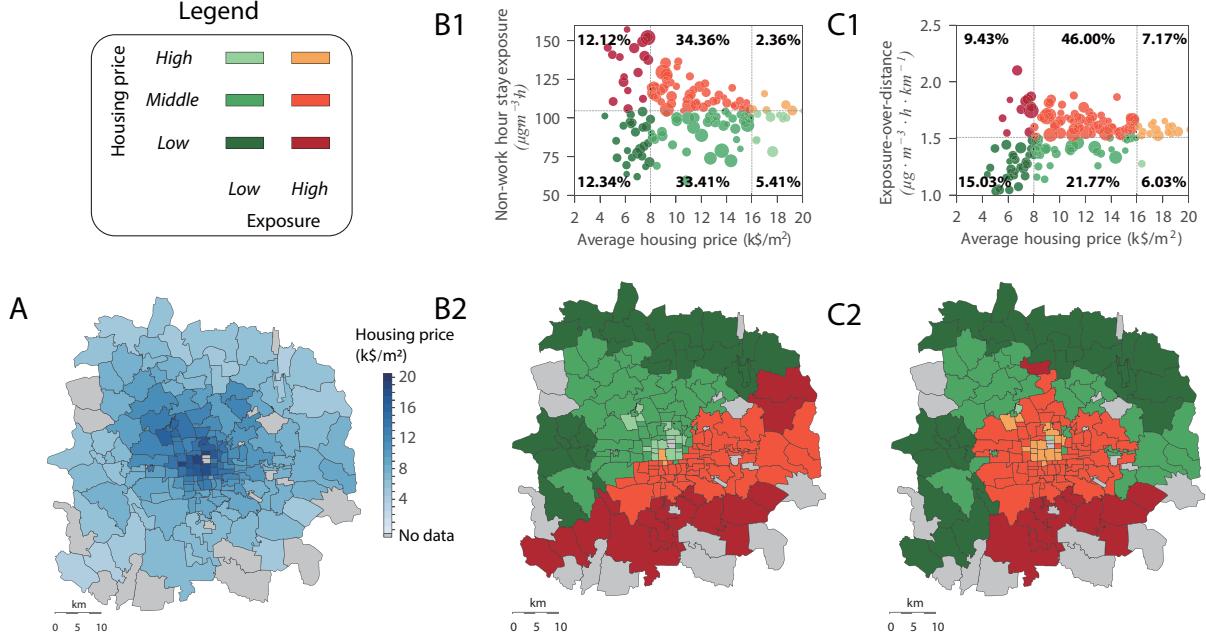
### 302 **3.3 Environmental justice in PM<sub>2.5</sub> stay- and travel-Exposure**

303 Environmental justice refers to “the fair treatment of all people with variant races, cul-  
304 tures and incomes, in development of regulations and policies.”<sup>29,30,53</sup> Here we investigate  
305 the environmental justice for commuters with different wealth levels, regarding their PM<sub>2.5</sub>  
306 exposure. We derive aggregated zonal housing price index from disaggregated housing price  
307 data, obtained from an online housing property listing platform in June 2016.<sup>54</sup> We use hous-  
308 ing price as a proxy for wealth level and examine its relationship with commuters’ hourly  
309 stay-exposure during non-work hours (which are mostly stay-at- home activities). We then  
310 compare their travel-exposures (EoD) across space.

311 Fig. 5A shows the average housing price in each zone, revealing higher housing value  
312 in the city center than in the suburbs. In each of the sub-figures from B1 to C2, the  
313 community zones in Beijing are separated into six groups by combining three levels of housing  
314 prices (i.e., low, middle and high) with two levels of PM<sub>2.5</sub> exposure (i.e., low and high).  
315 Fig. 5B1 and 5B2 display the relationship between the housing prices and the hourly stay  
316 exposure of commuters during the non-work hours in the winter, with the assumption that  
317 commuters go to work in the morning and return home in the evening. We estimate that  
318 hourly stay exposure to PM<sub>2.5</sub> for commuters with low, middle and high wealth levels are  
319  $111.82 \mu\text{g} \cdot \text{m}^{-3} \cdot \text{h}$ ,  $103.87 \mu\text{g} \cdot \text{m}^{-3} \cdot \text{h}$  and  $98.66 \mu\text{g} \cdot \text{m}^{-3} \cdot \text{h}$  on average in winter of 2015.  
320 That is, commuters with lower wealth are exposed to more PM<sub>2.5</sub> than their counterparts  
321 with higher wealth by 13% per hour when they stay at home. Moreover, the different groups  
322 of zones display clear differences. For example, 12.12% of the commuters in Beijing having  
323 high stay-exposure at home are with low level of wealth. Most of these population live in  
324 the southern suburbs, depicted in dark red; another 12.34% of the commuters are with low  
325 level of wealth but the PM<sub>2.5</sub> concentration in their residential areas were lower as they live

---

\*<http://www.mit.edu/~yanyanxu/exposure/>



**Figure 5: Relation between  $\text{PM}_{2.5}$  exposure and housing prices** **A** Average housing prices of each zone in thousand of US\$ per square meter ( $\text{k}\$ \cdot \text{m}^{-2}$ ) in June 2016. **B1, B2** Individual stay-exposure during non-work hours in winter of 2015 v.s. the average housing price in the same zone. Darker color refers to lower housing price and the green color refers to lower exposure. We classify zones into six groups based on housing price and exposure levels. **C1, C2** Exposure-over-distance (EoD) for commuters during the morning peak hour v.s. the average hosing price in the same zone.

326 in the north of the city. In short, for commuters in the southeastern Beijing, those with  
 327 lower wealth level experience higher level of  $\text{PM}_{2.5}$  exposure than those with higher wealth  
 328 levels. However, for commuters residing in the northwest of the city, those with lower wealth  
 329 level are exposed to less  $\text{PM}_{2.5}$ . This spatial disparity is mainly caused by the industrial and  
 330 economic activity distribution in the city and can be mitigated by future spatial planning  
 331 policy.

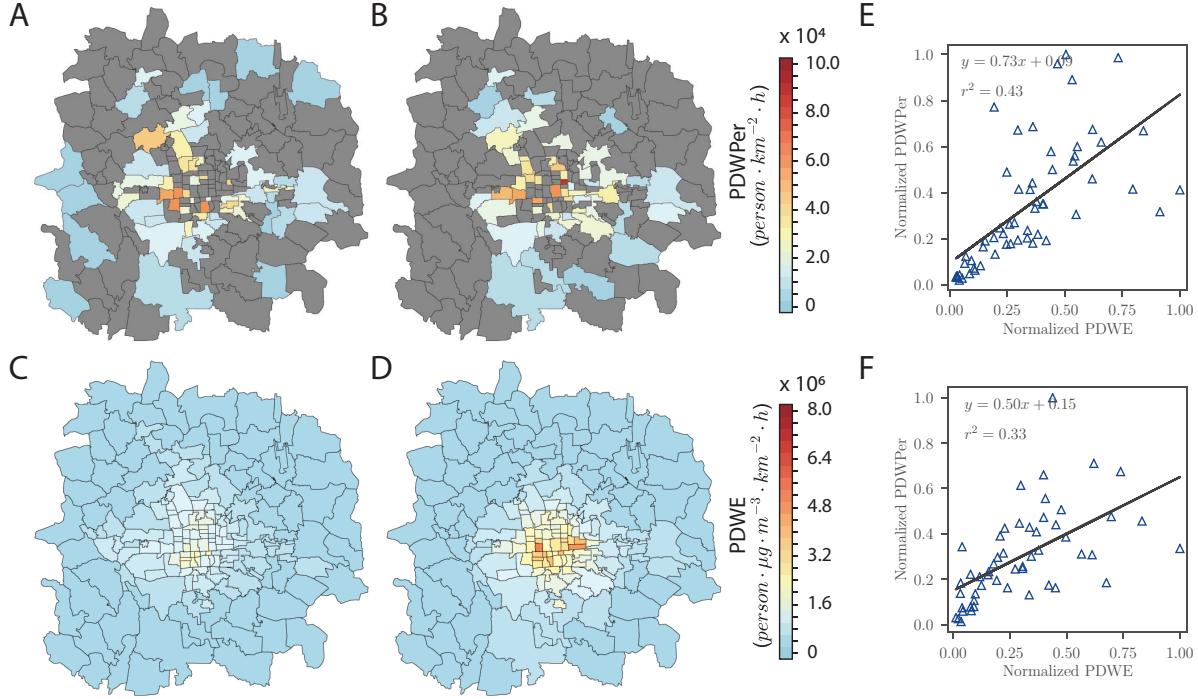
332 In contrast, the relationship between the housing price and the commuting travel exposure  
 333 (EoD) displays large spatial disparity, as shown in Fig. 5C1 and 5C2. The travel exposure  
 334 to  $\text{PM}_{2.5}$  per kilometer for commuters with low, middle and high wealth levels were  $1.47 \mu\text{g} \cdot$   
 $m^{-3} \cdot h \cdot \text{km}^{-1}$ ,  $1.56 \mu\text{g} \cdot m^{-3} \cdot h \cdot \text{km}^{-1}$  and  $1.55 \mu\text{g} \cdot m^{-3} \cdot h \cdot \text{km}^{-1}$  on average in winter of  
 335 2015. This indicates that the commuters with lower wealth were exposed to 5% less  $\text{PM}_{2.5}$   
 336 than commuters with higher wealth level when traveling the same distance for commuting  
 337

trips. The primary reason is that the lower wealth residents are living in the suburban areas where have lighter traffic congestion than the city center, as shown in Fig. S2A in the SM. We also estimate that 9.43% of the commuters had both low wealth and high travel exposure per kilometer during their commuting trips. They were concentrated in the southern areas, colored in dark red. For commuters residing in southern Beijing, those with low wealth level were exposed to more PM<sub>2.5</sub> for both stay-at-home activities and travels in the winter. Moreover, 46% of the commuters have middle wealth level and experienced high travel exposure (EoD) to PM<sub>2.5</sub> (as shown in the in orange zones). Their high EoD is mainly caused by the heavy traffic congestion within the 5th Ring Road in the south of Beijing.

The results from the summer are presented in Fig. S6 in the SM.

### 3.4 Perceived air quality experiment as a comparison

To compare the spatial correlation of PAQ and the objective estimates of PM<sub>2.5</sub> exposure, we then model the population density weighted air quality perception (PDWPer) with PAQ data on the same day (February 17th, 2016) by replacing the PM<sub>2.5</sub> concentration in PDWE with the average perception in each zone, similar to the calculation of PDWE with mobile phone data discussed previously. PDWPer is calculated only using the perception of participants and the static population density from census data. Only 97 individual samples for PAQ were completed for this day (February 17th, 2016) among the 256 individuals who responded for the 2 week study. The PDWPer at home and work are illustrated in Figs. 6A and 6B, respectively. As there is a very limited number of participants in the PAQ data, the PDWPer does not cover the Beijing area. Figs. 6C and 6D illustrate the PDWE of the worst hour during the non-work and work hours on the same day, respectively. Although the subjective perception of the participants does not contain the concentration of pollutants, we compare the exposures inferred from mobile phone data with PAQ data by normalizing both datasets. The  $r^2$  between exposure from mobile phone data and PAQ equals to 0.43 at home and 0.33 at work, as shown in Figs. 6E and 6F. Given the small sample size of PAQ and the



**Figure 6: Population exposure estimates from PAQ survey vs. mobile phone data.** **A, B** population density weighted air quality perception (PDWPer) at home and work per zone. The gray zones indicate that no participant lives or works in that zone. **C, D** population density weighted exposure (PDWE) during the worst non-work and work hour in each zone. **E, F** Normalized PDWPer and PDWE at home and work.

364 relative consistency between the two estimates from PAQ and mobile phone data, we show  
 365 that when survey data are expensive to collect, combining air quality monitoring of PM<sub>2.5</sub>  
 366 concentration, large-scale mobile phone data can be a good alternative to estimate exposure  
 367 to air pollution without surveys.

## 368 4. Discussion and conclusion

369 Exposure to air pollution threatens public health, increasing mortality and morbidity. Within  
 370 the same city, levels of exposure to air pollution differ in space and time. Among various  
 371 pollution metrics, PM<sub>2.5</sub> concentration is the major concern for the public in Beijing as it  
 372 is the main cause of haze and affects the heart and lungs when inhaled. On certain days,  
 373 schools need to be closed and people are encouraged to stay at home to avoid exposure to

<sup>374</sup> severe haze.

<sup>375</sup> Today, massive mobile phone data can help us better understand and simulate human  
<sup>376</sup> mobility at the metropolitan scale. Still, previous works that model population exposure  
<sup>377</sup> to PM<sub>2.5</sub> using mobile phone data only account for stays of the mobile phone users in each  
<sup>378</sup> zone<sup>13</sup> or estimate personal exposure to air pollutants using the dynamic stay locations  
<sup>379</sup> of individual mobile phone users.<sup>14,25</sup> The former work models the exposure of population  
<sup>380</sup> in each zone at aggregated level; the latter one focuses on the individual exposure and uses  
<sup>381</sup> mobile phone data to track the users. Both cases consider only the actual mobile phone users  
<sup>382</sup> instead of the total population and neglect the personal travel exposure, which is nearly 10%  
<sup>383</sup> of the total exposure if the resident spends 2 hours of travel per day. By introducing the  
<sup>384</sup> census data, we expand the mobility of mobile phone users to the population at scale and  
<sup>385</sup> their associated travel times. We estimate the PM<sub>2.5</sub> concentration in space and in the road  
<sup>386</sup> networks. Without having to rely on costly travel surveys, we can estimate for the entire  
<sup>387</sup> population their daily stay- and travel- exposure. In addition, we investigate environmental  
<sup>388</sup> justice regarding the relation between personal exposure to PM<sub>2.5</sub> and their level of wealth  
<sup>389</sup> using housing price as a proxy. We find that commuters residing in southern Beijing are both  
<sup>390</sup> economically disadvantaged and suffer both higher static and travel exposure in the winter.  
<sup>391</sup> This information is useful for policymakers to plan a more equitable city. Mitigation policies  
<sup>392</sup> may include subsidizing installation of air cleaners for low income population, regulating  
<sup>393</sup> heavily polluting factories, and planning for urban greening projects focusing on PM2.5  
<sup>394</sup> control.<sup>30,55</sup> Finally, we compare exposure during stays from two diverse data sources, one  
<sup>395</sup> passively collected via CDRs and one actively collected via a PAQ survey. We show that the  
<sup>396</sup> PM<sub>2.5</sub> exposure modeled by mobile phone data is able to corroborate what is obtained via  
<sup>397</sup> the survey.

<sup>398</sup> The mobile phone data used for the inference of urban mobility only cover about 0.5%  
<sup>399</sup> of the population in Beijing. Despite expanded with the actual population in census data,  
<sup>400</sup> such small sample size might cause bias in the estimation of travel demand at urban scale.

401 On the other hand, the low frequency of mobile phone usage may cause the loss of visited  
402 locations if the user doesn't interact with the cell phone in these places. However, these  
403 two shortcomings could be improved conveniently at low cost by expanding the sample of  
404 users and the duration of the datasets as the data of all mobile phone users have already  
405 been stored by telecommunication operators. Due to the lack of ground truth data (such  
406 as traffic counts or travel survey data), we are not able to directly validate our estimated  
407 travel demand in Beijing. The framework to generate the mobility model has been proved  
408 successful in some other cities, e.g., we have validated the mobility model in Boston with  
409 the United States National Household Travel Survey (NHTS) and the Massachusetts Travel  
410 Survey (MTS),<sup>40,41</sup> in Bay Area with the Bay Area Transportation Survey (BATS).<sup>50</sup> While  
411 mobile phone data is blind to the travel mode, in contrast with the Taxi GPS data and  
412 transit smart card data, it is still one of the best options to investigate the urban mobility  
413 in big cities, due to its high penetration rate.

414 It is noteworthy that the data sets used in this work were not collected for the same study  
415 time: The mobile phone data were acquired in 2013; the PM<sub>2.5</sub> concentrations from 2015  
416 to 2016; the housing price data from in 2016. However, as travel demand of population in  
417 metropolitan areas is stable at the aggregate level, traditional methods often use travel survey  
418 collected in every five to ten years to estimate travel demand in cities in most countries. For  
419 example, the national household travel survey are usually conducted in the US in every ten  
420 years. Therefore, the routine travel demand inferred from mobile phone data in 2013 can be  
421 reasonably well to estimate the stay and travel exposure of people in 2015. As we harness  
422 the housing price as a proxy for wealth distribution and the correlation between wealth level  
423 and housing price is relatively stable in big cities, the study of environmental justice will not  
424 be impacted by the short misalignment of data collection time which is limited by existing  
425 resources.

426 Future investigations in the following aspects would improve the estimation accuracy  
427 of population exposure to PM<sub>2.5</sub>: (1) modeling the infiltration for vehicles and buildings

<sup>428</sup> would improve the estimation accuracy of personal exposure. The PM<sub>2.5</sub> concentration  
<sup>429</sup> observed from the monitoring stations are adopted to model the population exposure without  
<sup>430</sup> accounting for the ambient PM<sub>2.5</sub> infiltration; (2) a chemistry-transport model would further  
<sup>431</sup> improve the estimate of PM<sub>2.5</sub> concentration and other pollutants in the road networks with  
<sup>432</sup> the consideration of land use and topography; (3) an individual mobility model with higher  
<sup>433</sup> resolution of data in the road segments would provide a more precise representation of  
<sup>434</sup> personal exposure.

## <sup>435</sup> Competing interests

<sup>436</sup> The author declares that he has no competing interests.

437 **References**

- 438 (1) Kampa, M.; Castanas, E. Human health effects of air pollution. *Environmental Pollu-*  
439 *tion* **2008**, *151*, 362–367.
- 440 (2) Pope III, C. A.; Ezzati, M.; Dockery, D. W. Fine-particulate air pollution and life  
441 expectancy in the United States. *New England Journal of Medicine* **2009**, 376–386.
- 442 (3) Lelieveld, J.; Evans, J.; Fnais, M.; Giannadaki, D.; Pozzer, A. The contribution of  
443 outdoor air pollution sources to premature mortality on a global scale. *Nature* **2015**,  
444 *525*, 367–371.
- 445 (4) World Health Organization and et al., *Ambient air pollution: a global assessment of*  
446 *exposure and burden of disease*; 2016.
- 447 (5) Kelly, F. J.; Zhu, T. Transport solutions for cleaner air. *Science* **2016**, *352*, 934–936.
- 448 (6) Beijing Municipal Environmental Monitoring Center. <http://www.bjmemc.com.cn/>,  
449 2016; [Online; accessed 15-September-2016].
- 450 (7) Di, Q.; Wang, Y.; Zanobetti, A.; Wang, Y.; Koutrakis, P.; Choirat, C.; Dominici, F.;  
451 Schwartz, J. D. Air pollution and mortality in the Medicare population. *New England*  
452 *Journal of Medicine* **2017**, *376*, 2513–2522.
- 453 (8) Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahsuvaroglu, T.;  
454 Morrison, J.; Giovis, C. A review and evaluation of intraurban air pollution exposure  
455 models. *Journal of Exposure Science and Environmental Epidemiology* **2005**, *15*, 185–  
456 204.
- 457 (9) Steinle, S.; Reis, S.; Sabel, C. E. Quantifying human exposure to air pollution-moving  
458 from static monitoring to spatio-temporally resolved personal exposure assessment.  
459 *Science of the Total Environment* **2013**, *443*, 184–193.

- 460 (10) Quiros, D. C.; Lee, E. S.; Wang, R.; Zhu, Y. Ultrafine particle exposures while walking,  
461 cycling, and driving along an urban residential roadway. *Atmospheric Environment*  
462 **2013**, *73*, 185–194.
- 463 (11) Levy, I.; Levin, N.; Schwartz, J. D.; Kark, J. D. Back-extrapolating a land use regres-  
464 sion model for estimating past exposures to traffic-related air pollution. *Environmental*  
465 *Science & Technology* **2015**, *49*, 3603–3610.
- 466 (12) Smith, J. D.; Mitsakou, C.; Kitwiroon, N.; Barratt, B. M.; Walton, H. A.; Taylor, J. G.;  
467 Anderson, H. R.; Kelly, F. J.; Beevers, S. D. London Hybrid exposure model: improv-  
468 ing human exposure estimates to NO<sub>2</sub> and PM<sub>2.5</sub> in an urban setting. *Environmental*  
469 *Science & Technology* **2016**, *50*, 11760–11768.
- 470 (13) Nyhan, M.; Grauwin, S.; Britter, R.; Misstear, B.; McNabola, A.; Laden, F.; Barrett, S.  
471 R. H.; Ratti, C. Exposure Track - The Impact of Mobile-Device-Based Mobility Pat-  
472 terns on Quantifying Population Exposure to Air Pollution. *Environmental Science &*  
473 *Technology* **2016**, *50*, 9671–9681.
- 474 (14) Dewulf, B.; Neutens, T.; Lefebvre, W.; Seynaeve, G.; Vanpoucke, C.; Beckx, C.; Van de  
475 Weghe, N. Dynamic assessment of exposure to air pollution using mobile phone data.  
476 *International Journal of Health Geographics* **2016**, *15*, 14.
- 477 (15) Shekarrizfard, M.; Faghih-Imani, A.; Tetreault, L.-F.; Yasmin, S.; Reynaud, F.;  
478 Morency, P.; Plante, C.; Drouin, L.; Smargiassi, A.; Eluru, N.; et al., Regional as-  
479 sessment of exposure to traffic-related air pollution: Impacts of individual mobility and  
480 transit investment scenarios. *Sustainable Cities and Society* **2017**, *29*, 68–76.
- 481 (16) Setton, E.; Marshall, J. D.; Brauer, M.; Lundquist, K. R.; Hystad, P.; Keller, P.;  
482 Cloutier-Fisher, D. The impact of daily mobility on exposure to traffic-related air pol-  
483 lution and health effect estimates. *Journal of Exposure Science and Environmental*  
484 *Epidemiology* **2011**, *21*, 42.



- 510 (25) Nyhan, M.; Kloog, I.; Britter, R.; Ratti, C.; Koutrakis, P. Quantifying population  
511 exposure to air pollution using individual mobility patterns inferred from mobile phone  
512 data. *Journal of exposure science & environmental epidemiology* **2018**,
- 513 (26) Gauderman, W. J.; Avol, E.; Gilliland, F.; Vora, H.; Thomas, D.; Berhane, K.; Mc-  
514 Connell, R.; Kuenzli, N.; Lurmann, F.; Rappaport, E.; et al., The effect of air pollution  
515 on lung development from 10 to 18 years of age. *New England Journal of Medicine*  
516 **2004**, *351*, 1057–1067.
- 517 (27) Devlin, R.; Ghio, A.; Kehrl, H.; Sanders, G.; Cascio, W. Elderly humans exposed  
518 to concentrated air pollution particles have decreased heart rate variability. *European  
519 Respiratory Journal* **2003**, *21*, 76s–80s.
- 520 (28) TomTom, TomTom Traffic Index, Measuring congestion worldwide. 2017; [https://  
521 www.tomtom.com/en\\_gb/trafficindex/](https://www.tomtom.com/en_gb/trafficindex/).
- 522 (29) Marmot, M. Social determinants of health inequalities. *The lancet* **2005**, *365*, 1099–  
523 1104.
- 524 (30) Brugge, D.; Patton, A. P.; Bob, A.; Reisner, E.; Lowe, L.; Bright, O.-J. M.; Du-  
525 rant, J. L.; Newman, J.; Zamore, W. Developing community-level policy and practice  
526 to reduce traffic-related air pollution exposure. *Environmental Justice* **2015**, *8*, 95–104.
- 527 (31) Gonzalez, M. C.; Hidalgo, C. A.; Barabasi, A.-L. Understanding individual human  
528 mobility patterns. *Nature* **2008**, *453*, 779–782.
- 529 (32) Jiang, S.; Yang, Y.; Gupta, S.; Veneziano, D.; Athavale, S.; González, M. C. The  
530 TimeGeo modeling framework for urban mobility without travel surveys. *Proceedings  
531 of the National Academy of Sciences* **2016**, 201524261.
- 532 (33) Jiang, S.; Ferreira, J.; González, M. C. Activity-based human mobility patterns inferred

- 533 from mobile phone data: A case study of Singapore. *IEEE Transactions on Big Data*  
534 **2017**, *3*, 208–219.
- 535 (34) Li, R.; Wang, W.; Di, Z. Effects of human dynamics on epidemic spreading in Côte  
536 d'Ivoire. *Physica A: Statistical Mechanics and its Applications* **2017**, *467*, 30–40.
- 537 (35) Guo, S.; Hu, M.; Zamora, M. L.; Peng, J.; Shang, D.; Zheng, J.; Du, Z.; Wu, Z.;  
538 Shao, M.; Zeng, L.; et al., Elucidating severe urban haze formation in China. *Proceedings  
539 of the National Academy of Sciences* **2014**, *111*, 17373–17378.
- 540 (36) Liu, J.; Mauzerall, D. L.; Chen, Q.; Zhang, Q.; Song, Y.; Peng, W.; Klimont, Z.;  
541 Qiu, X.; Zhang, S.; Hu, M.; et al., Air pollutant emissions from Chinese households:  
542 A major and underappreciated ambient pollution source. *Proceedings of the National  
543 Academy of Sciences* **2016**, *113*, 7756–7761.
- 544 (37) U.S. Environmental Protection Agency, Air Quality Designations for Particle Pollution.  
545 <https://www.epa.gov/particle-pollution-designations>, 2017; [Online; accessed 15-June-  
546 2017].
- 547 (38) Wong, D. W.; Yuan, L.; Perlin, S. A. Comparison of spatial interpolation methods  
548 for the estimation of air quality data. *Journal of Exposure Science and Environmental  
549 Epidemiology* **2004**, *14*, 404.
- 550 (39) Zou, B.; Luo, Y.; Wan, N.; Zheng, Z.; Sternberg, T.; Liao, Y. Performance comparison  
551 of LUR and OK in PM2. 5 concentration mapping: a multidimensional perspective.  
552 *Scientific Reports* **2015**, *5*.
- 553 (40) Alexander, L.; Jiang, S.; Murga, M.; González, M. C. Origin–destination trips by pur-  
554 pose and time of day inferred from mobile phone data. *Transportation research part c:  
555 emerging technologies* **2015**, *58*, 240–250.

- 556 (41) Çolak, S.; Alexander, L. P.; Alvim, B. G.; Mehndiratta, S. R.; González, M. C. An-  
557 alyzing cell phone location data for urban travel: current methods, limitations, and  
558 opportunities. *Transportation Research Record: Journal of the Transportation Research*  
559 *Board* **2015**, 126–135.
- 560 (42) Zheng, Y.; Xie, X. Learning travel recommendations from user-generated GPS traces.  
561 *ACM Transactions on Intelligent Systems and Technology (TIST)* **2011**, 2, 2.
- 562 (43) Jiang, S.; Fiore, G. A.; Yang, Y.; Ferreira Jr, J.; Frazzoli, E.; González, M. C. A  
563 review of urban computing for mobile phone traces: current methods, challenges and  
564 opportunities. Proceedings of the 2nd ACM SIGKDD International Workshop on Urban  
565 Computing. 2013; p 2.
- 566 (44) Guttman, A. *R-trees: a dynamic index structure for spatial searching*; ACM, 1984;  
567 Vol. 14.
- 568 (45) Beijing Regional Statistic Year Book. 2016; <http://www.bjstats.gov.cn/nj/qxnj/2014/zk/indexch.htm>, [Online; accessed 12-January-2016].
- 570 (46) Toole, J. L.; Colak, S.; Sturt, B.; Alexander, L. P.; Evsukoff, A.; González, M. C. The  
571 path most traveled: Travel demand estimation using big data resources. *Transportation*  
572 *Research Part C: Emerging Technologies* **2015**, 58, 162–177.
- 573 (47) Sass, V.; Kravitz-Wirtz, N.; Karceski, S.; Hajat, A.; Crowder, K.; Takeuchi, D. The  
574 effects of air pollution on individual psychological distress. *Health & place* **2017**, 48,  
575 72–79.
- 576 (48) Beijing Transportation Development & Research Center, The Third Beijing Travel  
577 Survey. <http://www.bjtrc.org.cn/>, 2007; [Online; accessed 12-March-2017].
- 578 (49) Huang, S.-s.; Song, R.; Tao, Y. Behavior of Urban Residents Travel Mode Choosing and

- 579 Influencing Factors - Taking Beijing as an Example. *Communications Standardization*  
580 **2008**, *9*, 1–5.
- 581 (50) Çolak, S.; Lima, A.; González, M. C. Understanding congested travel in urban areas.  
582 *Nature Communications* **2016**, *7*.
- 583 (51) OpenStreetMap. <https://www.openstreetmap.org>, 2016; [Online; accessed 18-April-  
584 2016].
- 585 (52) AMP direction API. [http://lbs.amap.com/api/webservice/reference/  
586 direction/](http://lbs.amap.com/api/webservice/reference/direction/), 2016; [Online; accessed 16-May-2016].
- 587 (53) Pearce, J.; Kingham, S.; Zawar-Reza, P. Every breath you take? Environmental justice  
588 and air pollution in Christchurch, New Zealand. *Environment and Planning A* **2006**,  
589 *38*, 919–938.
- 590 (54) Lianjia. <https://bj.lianjia.com>, 2016.
- 591 (55) Yang, J.; Chang, Y.; Yan, P. Ranking the suitability of common urban tree species for  
592 controlling PM2. 5 pollution. *Atmospheric Pollution Research* **2015**, *6*, 267–277.