# Supplementary Material for Universal Predictability of Mobility Patterns in Cities

Xiao-Yong Yan, Chen Zhao, Ying Fan, Zengru Di, and Wen-Xu Wang

## Contents

# S1 Applying the population-weighted opportunities model to six European cities

## S1.1 Data collection and preprocessing

We use Gowalla check-ins data set [S1] (http://snap.stanford.edu/data/loc-gowalla.html) to test the performance of our population-weighted opportunities (PWO) model in predicting the mobility patterns in European cities. Gowalla is a location-based social networking website in which users share their locations through checking-in. The data set includes in total 6,442,890 check-ins of users over the period of Feb. 2009 - Oct. 2010. For the data set we define a trip of a user's by his/her two consecutive check-ins at different locations. If a trip's origin and destination are both in same city, we classify the trip to inner-city trips. Based on the data, we sort out inner-city trips in six European cities that have a sufficient number of Gowalla users, including London (UK), Berlin (German), Prague (Czech), Oslo (Norway), Copenhagen (Denmark) and Goteborg (Sweden). Table S1 shows the number of trips in each city. According to the methods presented in section **Materials and methods, Data preprocessing** in the main text, we partition each city into equal-area square zones, each of which is of 1 $km^2$. Figure S1 shows the zone partition results and the population density distribution of the six cities.

Table S1: Data summary of six European cities.

| City | Number of Trips | Number of Zones |
|---|---|---|
| London | 49,323 | 986 |
| Berlin | 18,504 | 494 |
| Prague | 10,066 | 268 |
| Oslo | 11,464 | 88 |
| Copenhagen | 10,022 | 364 |
| Goteborg | 67,438 | 422 |

### S1.2 Prediction results

We compare the travel distance distributions and the travel fluxes between all pairs of locations produced by the PWO model and the radiation model. The results are respectively shown in figure S2 and figure S3, in which we can see that the predictions of the PWO model, including the distributions of travel distance and the travel fluxes between all pairs of locations are in good agreement with real observations, whereas the radiation model's results deviate from the real data.

## S2 Applying the PWO model to four additional U. S. cities

### S2.1 Data collection and preprocessing

We collected travel survey data from the *Metropolitan Travel Survey Archive* website (http://www.surveyarchive.org/), which records more than 40 U. S. cities' travel survey archives. Most surveys contain information of citizens, including their households, vehicles and a diary of their trips on a given day (including each trip's origin and destination location, start and end time, trip mode and purpose). Through checking the data we find that for most data sets the trip-endpoints are labelled by TAZ code or ZIP code. Due to the difficulty in converting the codes to geographic coordinates, we only use the survey data sets of four U. S. cities (New York, Seattle, Detroit and the Twin Cities) that contains the information of trip-endpoint's geographic coordinates (latitude and longitude).

We find out and record all trips in terms of the coordinates of their origins and destinations from the four data sets. Table S2 shows the number of trips and some other data descriptions of the four cities. Figure S4 shows the zone partition results and the population density distribution of the four cities.

### S2.2 Prediction results

Figure S5 shows the travel distance distributions produced by the PWO model as well as the radiation model. We can see that the PWO model can precisely reproduce the observed distributions of travel distance, whereas the prediction results of the radiation model deviate from the real data. We also compare the travel fluxes between all pairs

Table S2: Data summary of four U. S. cities.

| City | Survey Year | Households | Number of Trips | Number of Zones |
|---|---|---|---|---|
| New York | 1998 | 10,971 | 69,282 | 3,006 |
| Seattle | 2006 | 4,746 | 62,277 | 3,175 |
| Detroit | 1994 | 7,300 | 53,583 | 4,056 |
| Twin Cities | 2001 | 8,961 | 35,469 | 2,684 |

of locations produced by both models with the real data. As shown in Figure S6, we find that the average fluxes predicted by both models deviate from real observations to some extent. The prediction errors result from the low sampling rate in the household travel surveys. For instance, in New York, there are 3,006 zones and the travel matrix contains more than $9 \times 10^6$ elements in principle. However, the surveys only record 69,282 trips. In other word, more than 99% of the real travel matrix's elements are zeros. In contrast, the travel matrices established by the models are always fully filled (although the values of some elements are very small). Thus the fluxes predicted by both models inevitably deviate the insufficient samples of real fluxes. We believe adequate sampling rate of real fluxes will allow a fair comparison between reproduced results and real observations to validate our model.

## S3 Comparing the PWO model with parameterised models

Since the 1940s, many trip distribution models have been proposed for predicting human or freight mobility patterns. The gravity model [S2] and intervening opportunity model [S3] are two widely used models among them. Since both of them rely on specific parameters estimated in terms of real traffic data to predict mobility model, we name them parameterised models. A recently presented rank-based model [S4] also belongs to the parameterised models, although it needs very low input information to reproduce some key characteristics of human mobility patterns. In this section, we will compare the prediction performance of the PWO model with those of parameterised models.

### S3.1 The parameterised models

**(1) The gravity model**

The gravity model [S2] stems from Newton's gravity law and has many modified versions so far. The original gravity model has the form

$$T_{ij} = \alpha \frac{m_i m_j}{r_{ij}^{\beta}}, \tag{S1}$$

where $T_{ij}$ is the travels departed from location $i$ to location $j$, $m_i$ and $m_j$ are the populations of origin and destination and $r_{ij}$ is the distance between $i$ and $j$. Although this model has a very similar form with Newton's gravity law, the model's prediction results may violate the origin constraint $T_i = \sum_j T_{ij}$ and the destination constraint $T_j = \sum_i T_{ij}$. To ensure the constraints, one can alternatively use the doubly constrained gravity model [S5]

$$T_{ij} = A_i T_i B_j T_j f(r_{ij}), \tag{S2}$$

where $T_i$ is the total travels departed from location $i$, $T_j$ is the total travels arrived at location $j$, $f(r_{ij})$ is a function of the distance $r_{ij}$, and $A_i = 1/\sum_j B_j T_j f(r_{ij})$ as well as $B_j = 1/\sum_i A_i T_i f(r_{ij})$ are balancing factors that are interdependent to each other. An iterative process enables calculating $A_i$ and $B_j$, but it demands high computational complexity. To simplify the calculation, one can use the singly constrained versions, either origin or destination constrained, of the gravity model by setting one set of the balancing factors $A_i$ or $B_j$ equal to one.

Here we employ the origin-constrained gravity model [S5] to predict mobility patterns in cities, described as

$$T_{ij} = T_i \frac{m_j f(r_{ij})}{\sum_{k \neq i}^{N} m_k f(r_{ik})}. \tag{S3}$$

The distance function $f(r_{ij})$ can be of any forms, such as power or exponential function. Based on numerical test, we find that the gravity model with power function $f(r_{ij}) = r_{ij}^{-\beta}$ offers better characterization of the cities' mobility patterns than the exponential function (see figure S7 and figure S8). Thus we use the power distance function, the parameter $\beta$ of which is estimated by fitting the real travel data of eight cities (see details in **S2.2**).

**(2) The intervening opportunities (I. O.) model**

The I. O. model [S3] stresses that trip making is not directly related to distance but to the relative accessibility of opportunities for satisfying the objective of the trip. The model's basic assumption is that for every trip departed from a location, there is a constant probability $p$ that determines a traveller being satisfied with a single opportunity. If a location $j$ has $m_j\Theta$ opportunities (we assume the number of opportunities at a location $j$ is proportional to its population $m_j$), the probability of a traveller being attracted by location $j$ is $\alpha m_j$, where $\alpha = p\Theta$.

Considering now the probability $q_i^j$ of not being satisfied by any of the opportunities offered by the $jth$ destinations away from the origin $i$, we can write

$$q_i^j = q_i^{j-1}(1 - \alpha m_j) \tag{S4}$$

or

$$\frac{q_i^j - q_i^{j-1}}{q_i^{j-1}} = -\alpha m_j = -\alpha(S_{i,j} - S_{i,j-1}), \tag{S5}$$

where $S_{ij}$ is the total population between location $i$ and $j$ (including $i$ and $j$). Assuming that the number of destinations is sufficiently large, we can treat $q$ and $S$ as continuous variables. Then Eq. (S5) can be rewritten as

$$\frac{\mathrm{d}q_i}{q_i(S)} = -\alpha \mathrm{d}S. \tag{S6}$$

After integration we obtain

$$q_i(s) = \frac{e^{-\alpha S}}{1 - e^{-\alpha M}}, \tag{S7}$$

where $M$ is the total population in the city. Note that the trip departed from location $i$ to location $j$ is equal to

$$T_{ij} = T_i[q_i(S_{i,j-1}) - q_i(S_{ij})]. \tag{S8}$$

Combining Eq. (S8) and Eq. (S7), we obtain the I. O. model:

$$T_{ij} = T_i \frac{e^{-\alpha(S_{ij} - m_j)} - e^{-\alpha S_{ij}}}{1 - e^{-\alpha M}}. \tag{S9}$$

**(3) The rank-based model**

The rank-based model [S4] assumes that the probability of an individual travelling from an origin to a destination depends (inversely) only upon the rank-distance between

the destination and the origin. The model is described as

$$T_{ij} = T_i \frac{R_i(j)^{-\gamma}}{\sum_{k \neq i}^{N} R_i(j)^{-\gamma}}, \tag{S10}$$

where $R_i(j)$ is the rank-distance from location $j$ to $i$ (*e.g.*, if $j$ is the closest location to $i$, $R_i(j) = 1$; if $j$ is the second closest location to $i$, $R_i(j) = 2$) and $\gamma$ is an adjustable parameter.

### S3.2 Estimating model parameters

Before applying the parameterised models, it is necessary to estimate their parameters. The goal of the parameter estimation is to maximise the accuracy of reproducing real mobility patterns by the models. Here we use Hyman method [S6], a standard method for calibrating gravity model in transportation planning [S5], to identify the gravity model's parameter.

Hyman method aims to find an optimal parameter to minimise the difference between modelled average travel distance and real average travel distance

$$E(\beta) = |\bar{r}(\beta) - \bar{r}| = \left| \frac{\sum_i \sum_j T_{ij}(\beta) r_{ij}}{\sum_i \sum_j T_{ij}(\beta)} - \frac{\sum_i \sum_j T_{ij} r_{ij}}{\sum_i \sum_j T_{ij}} \right|, \tag{S11}$$

where $\bar{r}(\beta)$ is the average distance given by the gravity model with parameter $\beta$, $\bar{r}$ is the real average travel distance, $T_{ij}(\beta)$ is the number of travels from zone $i$ to $j$ generated by the gravity model and $T_{ij}$ is the real number of travels from zone $i$ to $j$. It is not easy to solve the equation $E(\beta) = 0$. Hyman suggests that we use the secant method to address this problem, described by the following process:

**Step 1**. Give an initial estimate of $\beta_0 = 1/\bar{r}$.

**Step 2**. Calculate a trip matrix using the gravity model with the parameter $\beta_0$ and obtain a modelled average travel distance $\bar{r}(\beta_0)$. Estimate a better value of $\beta$ by means of

$$\beta_1 = \beta_0 \bar{r}(\beta_0)/\bar{r}. \tag{S12}$$

**Step 3**. Applying the gravity model with the estimated value of $\beta$ to calculate a new trip matrix and obtain a newly modelled average travel distance $\bar{r}(\beta)$ to compare with $\bar{r}$ . If they are sufficiently close to each other, terminate the iteration and accept the newest value of $\beta$ as the best estimation; otherwise go to step 4.

**Step 4**. Improve the estimation of $\beta$ via:

$$\beta_{i+1} = \frac{(\bar{r} - \bar{r}(\beta_{i-1}))\beta_i - (\bar{r} - \bar{r}(\beta_i))\beta_{i-1}}{\bar{r}(\beta_i) - \bar{r}(\beta_{i-1})}. \tag{S13}$$

**step 5**. Repeat steps 3 and 4 until $\bar{r}(\beta)$ is sufficiently close to $\bar{r}$.

The estimated parameters of the gravity model, the I.O. model and the rank-based model by Hyman method are listed in Table S2 for different cities.

Table S3: Estimation of parameter values in three parameterised models

| City | Gravity model | I. O. model | Rank-based model | Avg. travel distance |
|---|---|---|---|---|
| | $\beta$ | $\alpha$ | $\gamma$ | $\bar{r}$ (km) |
| Beijing | 1.71 | $5.71\times10^{-6}$ | 1.14 | 5.50 |
| Shenzhen | 1.63 | $3.41\times10^{-6}$ | 1.19 | 4.77 |
| Abidjan | 2.43 | $3.04\times10^{-5}$ | 1.36 | 2.47 |
| Chicago | 2.14 | $2.69\times10^{-4}$ | 1.15 | 9.49 |
| New York | 2.28 | $5.45\times10^{-4}$ | 1.20 | 11.71 |
| Seattle | 1.93 | $2.65\times10^{-4}$ | 1.12 | 8.30 |
| Detroit | 2.02 | $4.59\times10^{-4}$ | 1.13 | 8.94 |
| Twin Cities | 1.93 | $3.56\times10^{-4}$ | 1.04 | 8.15 |

### S3.3 Comparison among different models

**(1) Travel distance distribution**

Figure S9 shows the travel distance distribution predicted by different models. We see that both the gravity model and the rank-based model can reproduce the observed distributions of travel distance in most cases, but the I. O. model's results have significant deviation from the real data. Although in some cases the results of the gravity model (or rank-based model) are better than those of the PWO model, the gravity model needs a distance-decay function with adjustable parameters to match real data, whereas our model relies solely on the population distribution without free parameters.

**(2) Destination travel constraints**

Figure S10 shows the destination travel constraints produced by the models. We see that the results from the PWO model are in equal or better agreement with the real data than those of the other models in all cases. It is noteworthy that the rank-based

model shows the worst performance in this aspect. We speculate that the fact that the rank-based model do not make use of the population data at destination locations accounts for the big difference from real observations.

**(3) Travel fluxes between all pairs of locations**

We compare the travel fluxes between all pairs of locations predicted by models with empirical data. As shown in figure S11 and figure S12, we observe that all the predicted average fluxes by the PWO model, the gravity model and the rank-based model are comparable with the real fluxes to some extent. To give a more explicit comparison among different models, we use Sørensen similarity index as an alternative to measure the degree of agreement between reproduced travel matrices and empirical observations. Figure S13 shows that on average, the accuracy of the PWO model is higher than that of the I. O. model and the rank-based model. Although in some cases the gravity model can yield better prediction accuracy than the PWO model, the gravity model needs parameter values estimated from previous mobility measurements. In contrast, the PWO model only require the population distribution as input, rendering its application scope broader.

## S4 Relationship among trip distribution models

To deepen our understanding of the underlying mechanism in the trip distribution models explored in this paper, we discuss the relationship among them. We will first show that, in the particular case of uniform population distribution, the PWO, the radiation model, the I. O. model and the rank-based model can all transform into gravity-like models. Next, we will show that these models can be classified into two categories of modelling frameworks: sequential selection and global selection.

### S4.1 Uniform population distribution

Consider a uniform population distribution (*i.e.* $S_{ji} = \rho\pi r_{ij}^2$, where $\rho$ is the population density). We can write the PWO model (Eq. (1) in the main text) as

$$T_{ij} = T_i \frac{m_j(r_{ij}^{-2} - \frac{\pi}{A})}{\sum_{k\neq i}^{N} m_k(r_{ij}^{-2} - \frac{\pi}{A})}, \tag{S14}$$

where $A$ is the area of the city. Comparing with Eq. (S3), we can realise that Eq. (S14) is actually a gravity model with the distance function $f(r_{ij}) = r_{ij}^{-2} - \frac{\pi}{A}$. This function is a power law with a cut-off (see figure S14(A)). Since the population is not uniformly distributed in real cities, we can not directly use such distance function in the gravity model to predict travel fluxes. Alternatively, we have to estimate its parameters by relying on real traffic data prior to applying the gravity model. However, we may directly choose a population function, such as $f(S_{ji}) = \frac{1}{S_{ji}} - \frac{1}{M}$, to be used in the PWO model in the sense that the heterogeneity of population distribution has been captured by $S_{ji}$. Figure S14(B) shows the relationship between the population $S_{ji}$ and the travel proportion $T_{ij}/T_i$ in Abidjan, which is in agreement with the population function.

When the population distribution is of uniform distribution ($s_{ij} = \rho\pi r_{ij}^2$), the radiation model [S7] is reduced to

$$T_{ij} \propto T_i m_j r_{ij}^{-4}, \tag{S15}$$

which is actually a gravity model with a power-law distance function with power exponent $\beta = 4$. Comparing with the uniform version of the PWO model (having a power exponent $\beta = 2$), the selection scope of an individual in the radiation model is relatively more local. The radiation model can characterise the mobility patterns at the country scale (the estimated power exponent of gravity law in the case of U. S. state-wide commuting trips is 3.05 [S7], which does not significantly deviate from 4), but it is not applicable to predicting city mobility patterns. Table S2 shows that the estimated power exponents of the cities are subject to the range $1.63 - 2.43$, quite close to the exponent in Eq. (S14) but different from that in Eq. (S15).

The I. O. model can be also transformed into a gravity-like form in the case of uniform population distribution:

$$T_{ij} \propto T_i(e^{\alpha m_j} - 1)e^{-\alpha S_{ij}} = T_i(e^{\alpha m_j} - 1)e^{-\lambda r_{ij}^2}, \tag{S16}$$

where $\lambda = \alpha\rho\pi$. The distance function is of high-order exponential form, implying the lack of long-distance travel generated by the model. Thus it is not surprise that the I. O. model usually underestimates long-distance travels, as shown in figure S9.

The rank-based model uses rank-distance rather than spatial distance to predict the travels between locations. When the population are uniformly distributed in cities, the

rank-distance between locations is proportional to the square of the spatial distance, such that the rank-based model can be rewritten as

$$T_{ij} \propto T_i r_{ij}^{-2\gamma}. \tag{S17}$$

The distance function in Eq. (S17) is a power law with the power exponent around 2 (see Table S2). It can thus yield similar results to the travel distance distribution resulting from the gravity model and the PWO model (see figure S9). However, in the rank-based model the information of population in destination is ignored, rendering the destination travel constraints inaccurate, as shown in figure S10.

Taken together, insofar as given uniform population distributions, the PWO model, the radiation model, the I. O. model and the rank-based model can presented to be gravity-like models. Although these models have different hypothesis, they share similar underlying mechanism: the probability that an individual selects a travel destination is decreased along with the increment of some prohibitive factors. In the gravity models, the factor is spatial distance; in the rank-based model it is the rank-distance; in the I. O. model, the radiation model or the PWO model, it is the population between origins and destinations. The key difference lies in the fact that the gravity model, the I. O. model and the rank-based model need adjustable parameters to quantify the decrement effect, whereas in the radiation model and the PWO model, the decrement effect is naturally determined by population distribution.

### S4.2 Sequential selection and global selection

According to the decision-making process of travellers for selecting destinations, the frameworks of predicting mobility patterns can be classified into two categories. The first category includes the I. O. model and the radiation model, in which each traveller ranks potential destinations in ascending order according to the distance to his/her origin. An individual first decides whether to travel to the first destination in terms of a probability which is determined by some specific rules (see schematic in figure S15(B)). If the individual abandons the destination, the second one will be considered in terms of the same probability. Analogously, all potential destinations will be considered step by step until the individual eventually decides to travel to a chosen one. We name such step-by-step decision-making process *sequential selection*. The modelling framework

can be described in a unified form:

$$q_i^j = q_i^{j-1}(1-\theta_j), \tag{S18}$$

where $q_i^j$ is the probability of excluding the 1st to $j$th destinations departed from the origin $i$ and $\theta_j$ is the probability of selecting $j$th destination insofar as the 1st to $(j-1)$th destinations are not selected. Thus, the probability of selecting $j$ to travel can be expressed in a joint probability

$$p_{ij} = q_i^{j-1}\theta_j. \tag{S19}$$

The I. O. model can be derived by assuming that the probability $\theta_j$ is proportional to the population of destination $j$ (*i.e.* $\theta_j = \alpha m_j$). After some calculations (see details in section **S2.1**), we finally have

$$p_{ij} = \frac{e^{-\alpha(S_{ij}-m_j)} - e^{-\alpha S_{ij}}}{1-e^{-\alpha M}}, \tag{S20}$$

which is the probability of selecting destination $j$ departed from $i$ in the I. O. model.

Similarly, assuming that the probability $\theta_j$ is the ratio of the population of destination $j$ to the total population $S_{ij}$ between locations $i$ and $j$, we have

$$q_i^j = q_i^{j-1}(1-m_j/S_{ij}) = \prod_j \frac{S_{i,j-1}}{S_{ij}} = \frac{m_i}{S_{ij}} \tag{S21}$$

and

$$p_{ij} = q_i^{j-1}\frac{m_j}{S_{ij}} = \frac{m_i m_j}{S_{i,j-1}S_{ij}}, \tag{S22}$$

which is nothing but the radiation model.

Different forms of probability $\theta_j$ can lead to different versions of models that are subject to the framework of sequential selection. For instance, assuming the probability $\theta_j$ is the ratio of $m_j$ to the remaining population $M - S_{ij} + m_j$, we can have

$$q_i^j = q_i^{j-1}(1-\frac{m_j}{M-S_{ij}+m_j}) = \prod_j \frac{M-S_{ij}}{M-S_{ij}+m_j} = \frac{M-S_{ij}}{M} \tag{S23}$$

and

$$p_{ij} = q_i^{j-1}\frac{m_j}{M-S_{ij}+m_j} = \frac{m_j}{M}, \tag{S24}$$

which is the uniform selection model [S8].

Note that in the sequential selection models, it is possible to find that travellers do not select any destinations to travel unless the system is infinite [S9]. In general, the probability is $p_{ii} = 1 - \sum_j p_{ij}$. In other words, a traveller stays at the origin with probability $p_{ii}$. For the I. O. model, $p_{ii} = (1 - e^{-\alpha m_i})/(1 - e^{-\alpha M})$; for the radiation model and the uniform selection model, $p_{ii} = m_i/M$.

The second category, named *global selection*, includes the gravity model, the PWO model and the rank-based model. In global selection model's decision-making process, a traveller evaluates the attractions of all possible destinations simultaneously and selects a destination to travel with a probability proportional to the destination's attraction (see schematic in figure S15(C)). The unified framework can be described as

$$p_{ij} = \frac{A_j}{\sum_j A_j}, \tag{S25}$$

where $A_j$ is the attraction of destination $j$. If we solely use population to capture the destination's attraction, the uniform selection model is obtained; if we use some functions to describe the decay of attraction along with the increase of (real or rank) distance, we can obtain the gravity or the rank-based model (see Eq. (S3) and Eq. (S10)); if the attraction is inversely proportional to the population between destinations and origins, the PWO model is derived.

Despite the difference between the two modelling frameworks, both sequential and global selection models imply the preference for closer destinations in human travel decision-making: in sequential selection model, a closer destination has a higher priority to be selected; in global selection model, the attraction of a closer destination decays slowly than that of a farther one. Although both frameworks capture the decision-making process of travellers to some extent, our comparison study (figure S7-S11) demonstrates that at the city scale, global selection models perform better than sequential selection models.

## References

[S1] Cho E, S. Myers A, Leskovec J. Friendship and mobility: user movement in location-based social networks. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2011.

[S2] Zipf GK. 1946 The $P_1P_2/D$ hypothesis: on the intercity movement of persons. *Am. Sociol. Rev.* **11**, 677-686.

[S3] Stouffer SA. 1940 Intervening opportunities: a theory relating mobility and distance. *Am. Sociol. Rev.* **5**, 845-867.

[S4] Noulas A, Scellato S, Lambiotte R, Pontil M, Mascolo C. 2012 A tale of many cities: universal patterns in human urban mobility. *PLoS ONE* **7**, e37027.

[S5] Ortúzar JD, Willumsen LG. 2011 *Modelling Transport.* New York, USA: John Wiley & Sons.

[S6] Hyman GM. 1969 The calibration of trip distribution models. *Env. Plan.* **1**, 105-112.

[S7] Simini F, González MC, Maritan A, Barabási A-L. 2012 A universal model for mobility and migration patterns. *Nature* **484**, 96-100.

[S8] Simini F, Maritan A, Néda Z. 2013 Human mobility in a continuum approach. *PLoS ONE* **8**, e60069.

[S9] Masucci AP, Serras J, Johansson A, Batty M. 2013 Gravity versus radiation model: on the importance of scale and heterogeneity in commuting flows. *Phys. Rev. E* **88**, 022812.

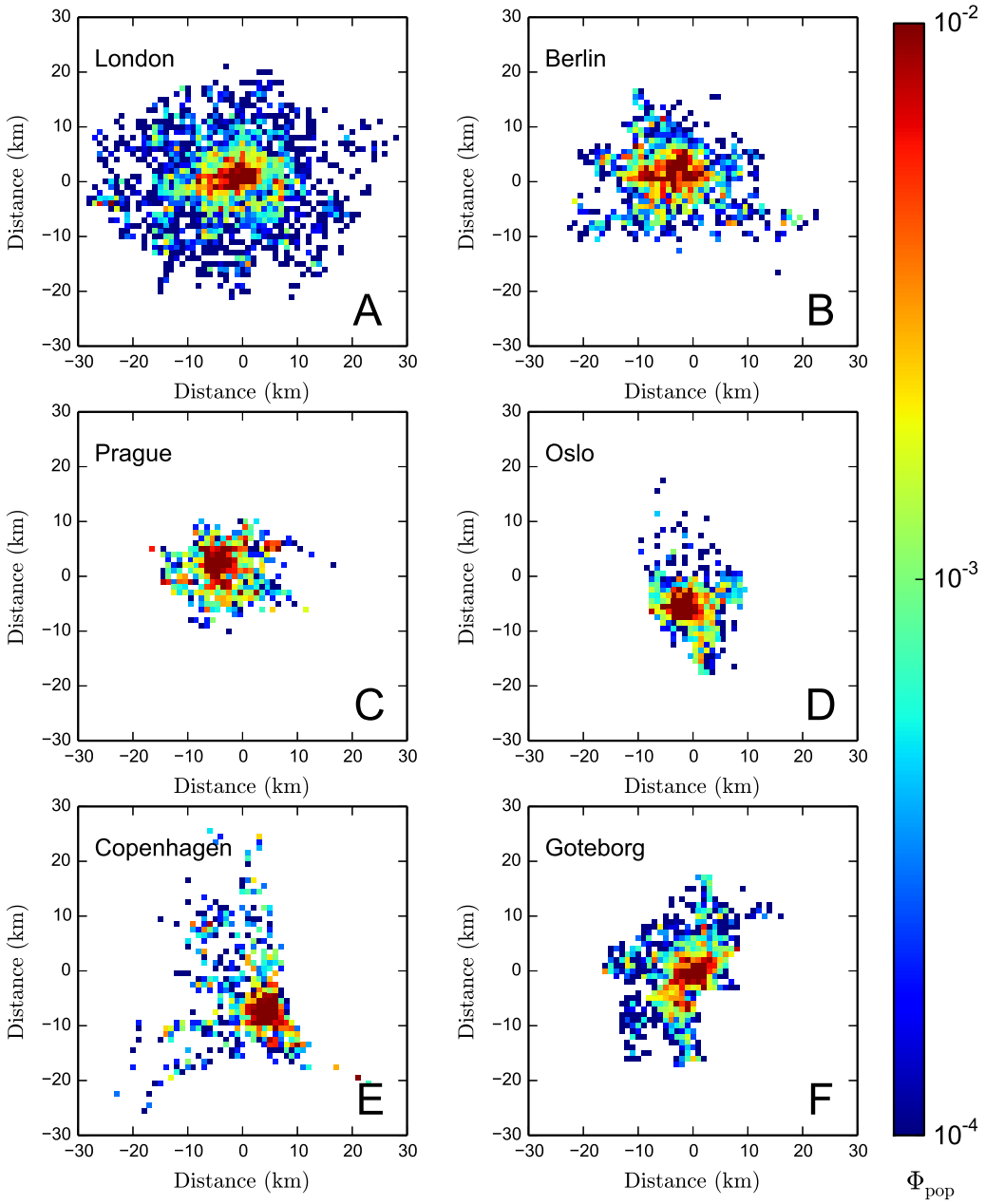


Figure S1: **The zone partition and population density distribution of six European cities**. (A) London. (B) Berlin. (C) Prague. (D) Oslo. (E) Copenhagen. (F) Goteborg. The density function $\Phi_{pop}(i)$ represents the probability of finding a travel started from zone $i$.

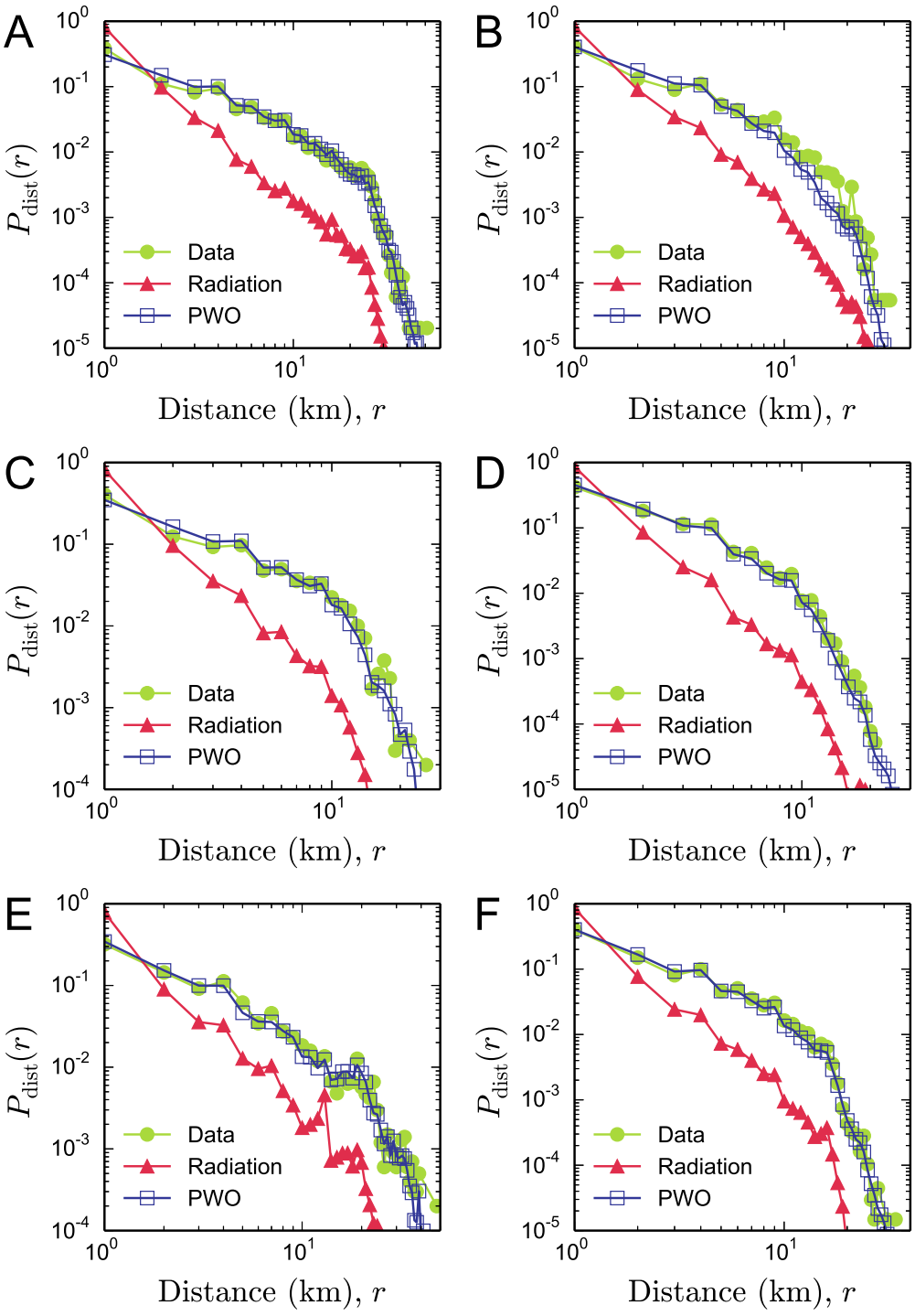


Figure S2: **Comparing the travel distance distributions generated by different models**. (A) London. (B) Berlin. (C) Prague. (D) Oslo. (E) Copenhagen. (F) Goteborg. $P_{\mathrm{dist}}(r)$ is the probability of a travel between locations at distance $r$.

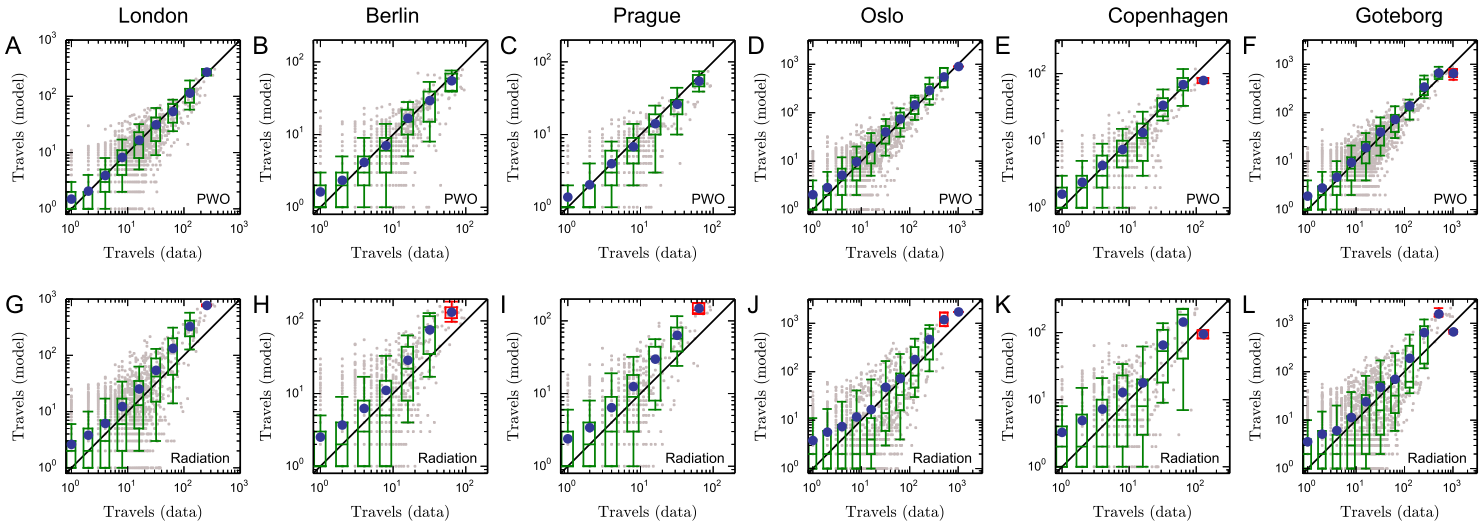


Figure S3: **Comparing the observed fluxes with the predicted fluxes**. (A-F) Predicted fluxes generated by the PWO model. (G-L) Predicted fluxes generated by the radiation model.

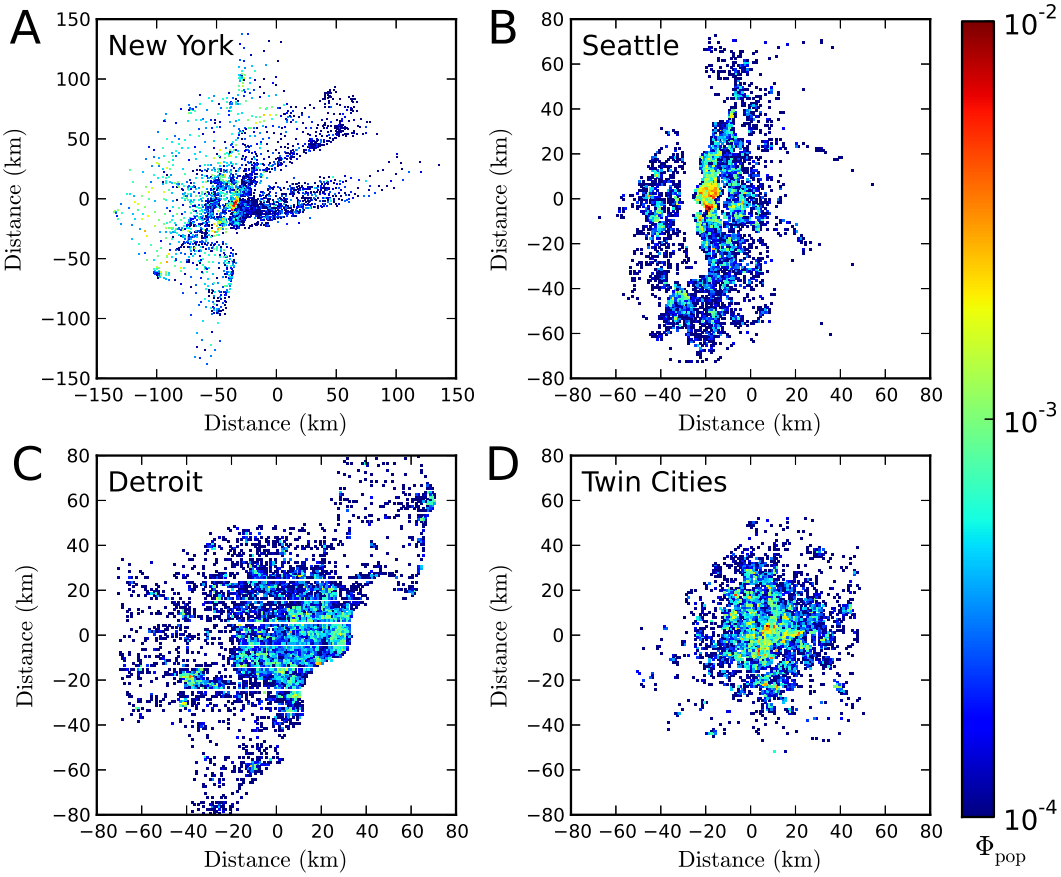


Figure S4: **The zone partition and population density distribution of four U. S. cities**. (A) New York. (B) Seattle. (C) Detroit. (D) Twin Cities. The density function $\Phi_{pop}(i)$ represents the probability of finding a travel started from zone $i$.

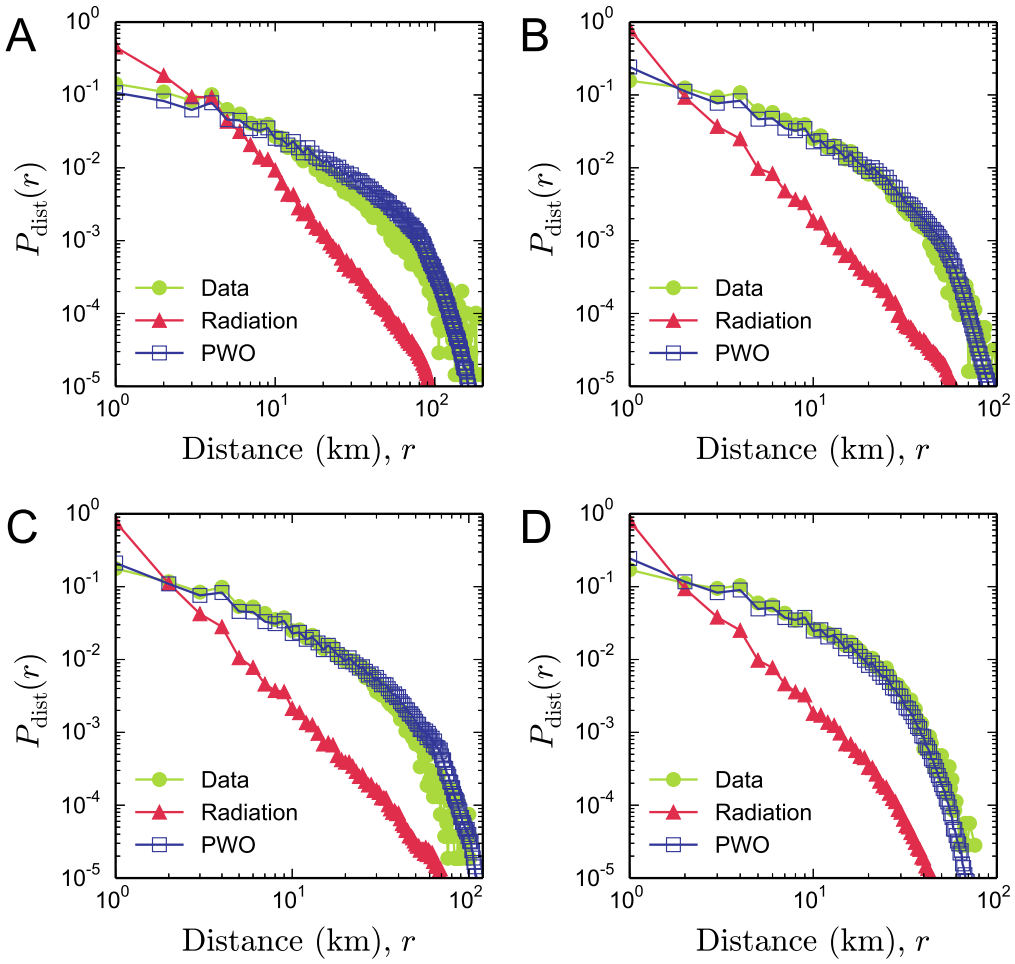


Figure S5: **Comparing the travel distance distributions generated by different models**. (A) New York. (B) Seattle. (C) Detroit. (D) Twin Cities. $P_{\mathrm{dist}}(r)$ is the probability of a travel between locations at distance $r$.

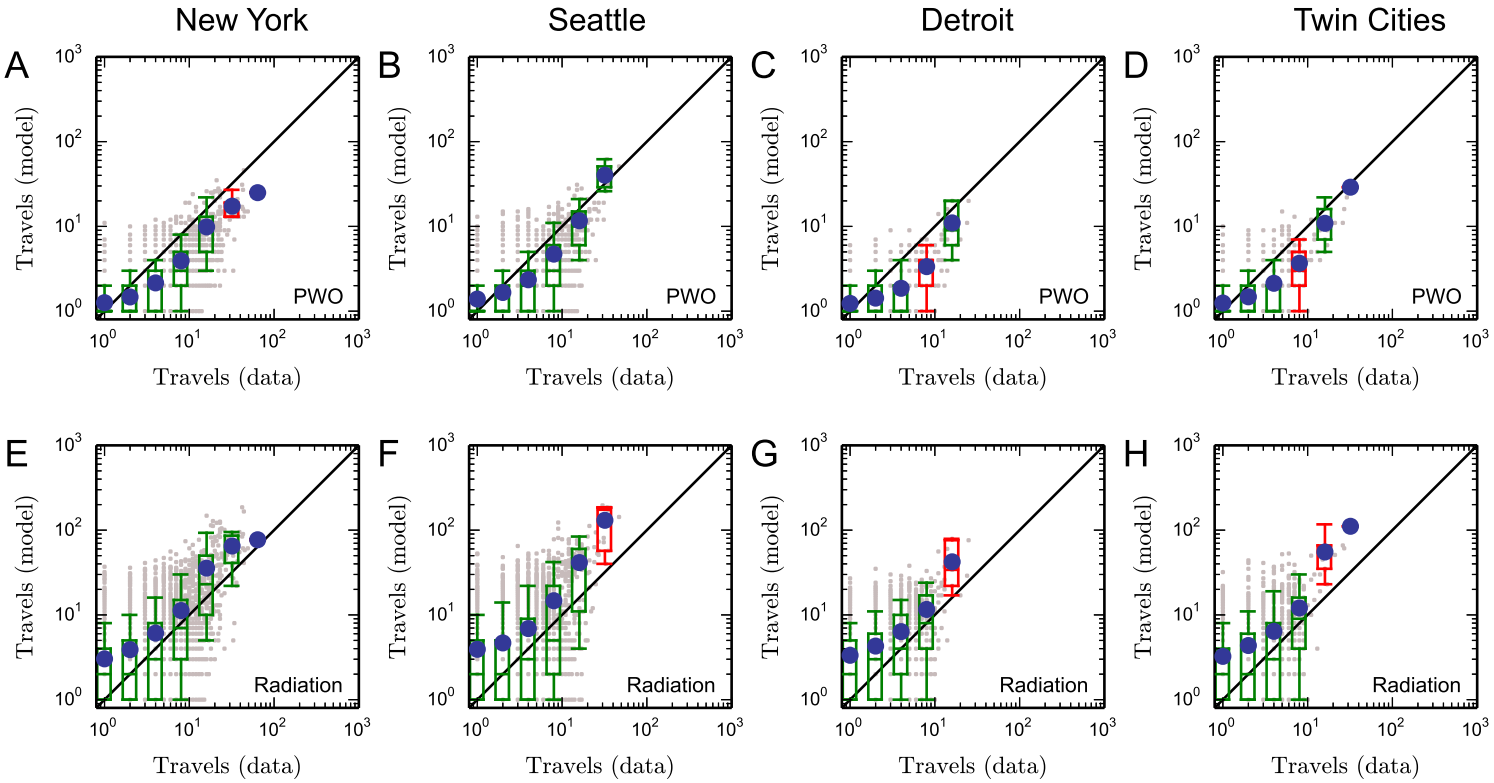


Figure S6: **Comparing the observed fluxes with the predicted fluxes**. (A-D) Predicted fluxes generated by the PWO model. (E-H) Predicted fluxes generated by the radiation model.

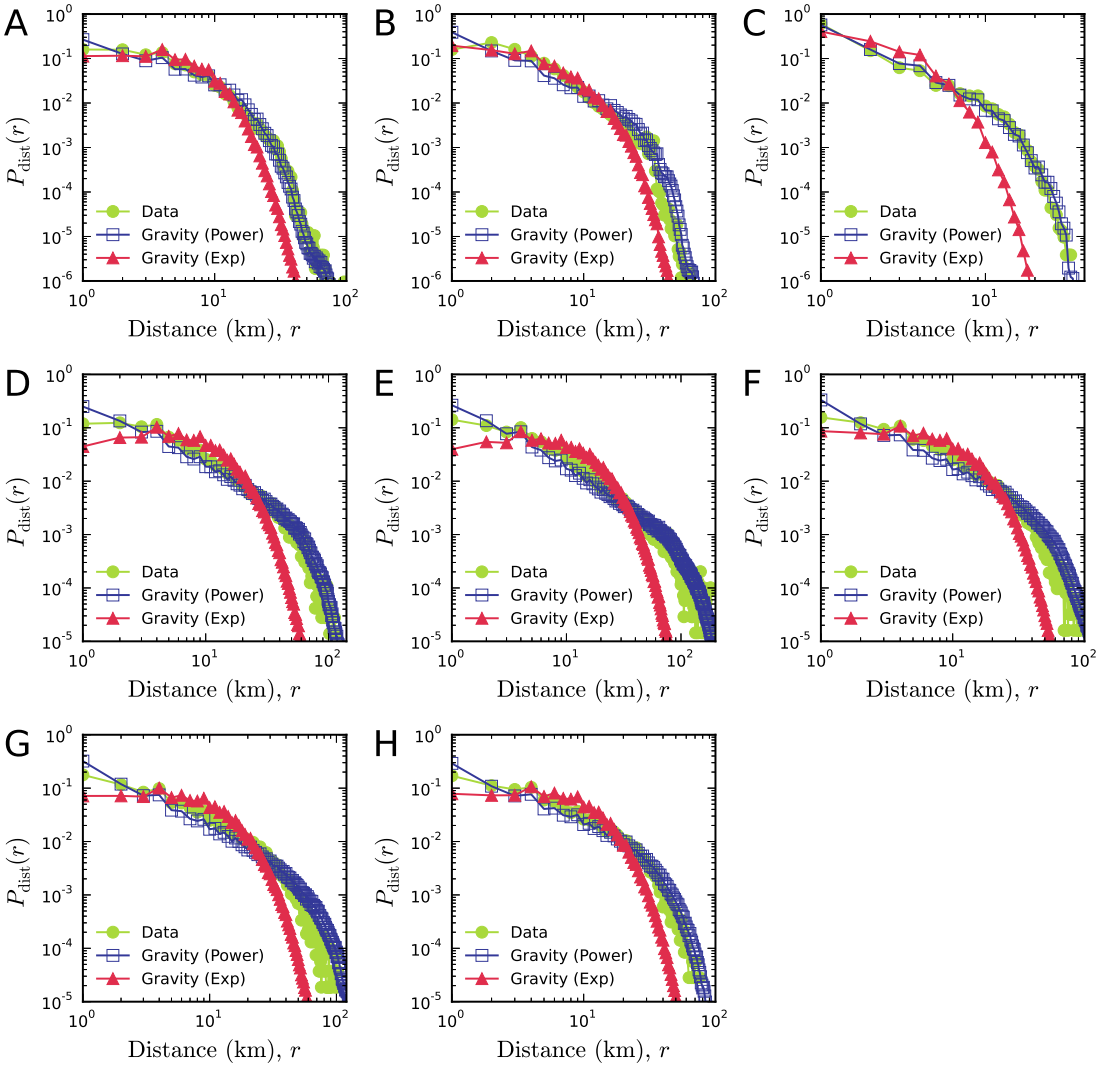


Figure S7: **Comparing the travel distance distributions generated by two types of gravity models**. (A) Beijing. (B) Shenzhen. (C) Abidjan. (D) Chicago. (E) New York. (F) Seattle. (G) Detroit. (H) Twin Cities.

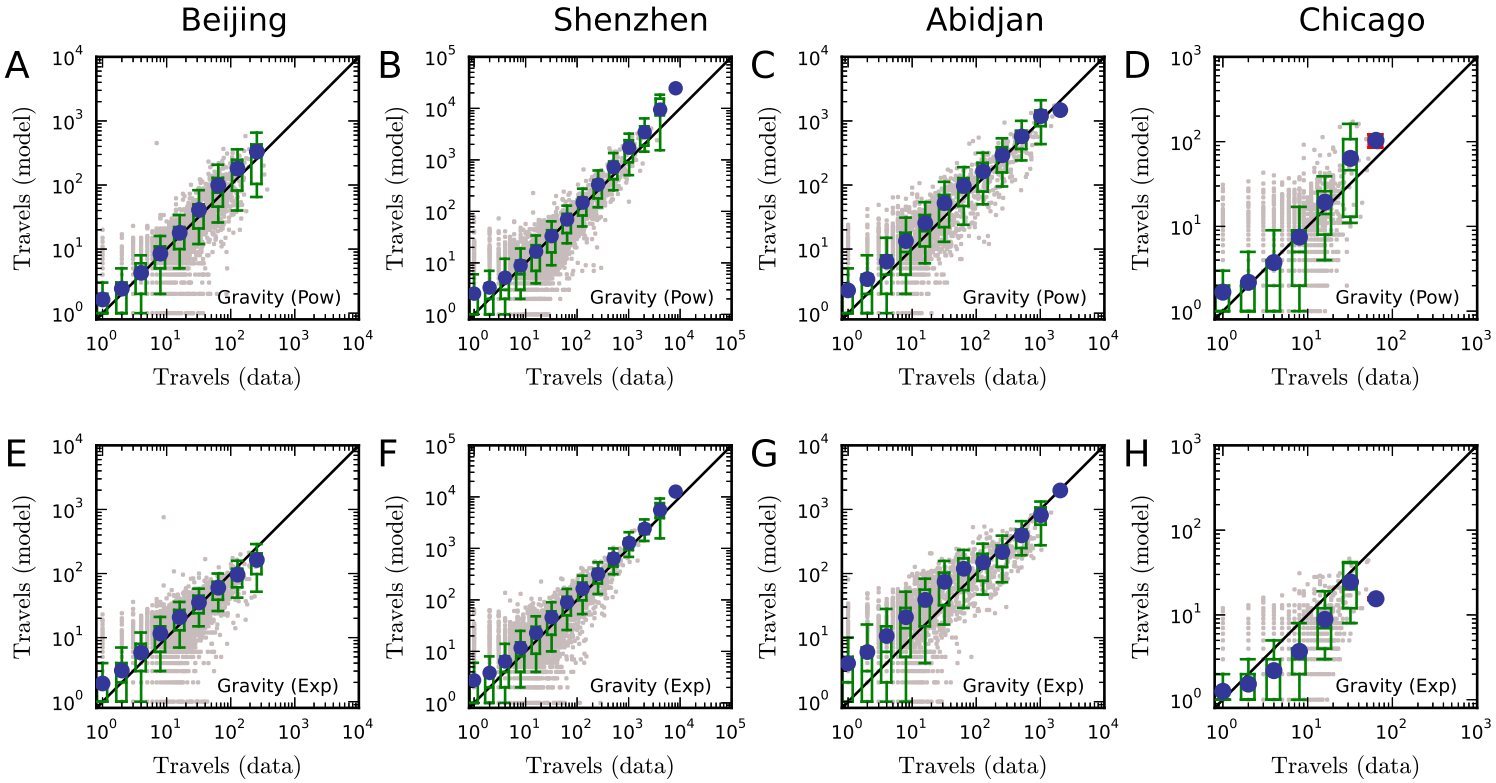


Figure S8: **Comparing the observed fluxes with the predicted fluxes**. (A-D) Predicted fluxes generated by the gravity model with power-law distance function. (E-H) Predicted fluxes generated by the gravity model with exponential distance function.

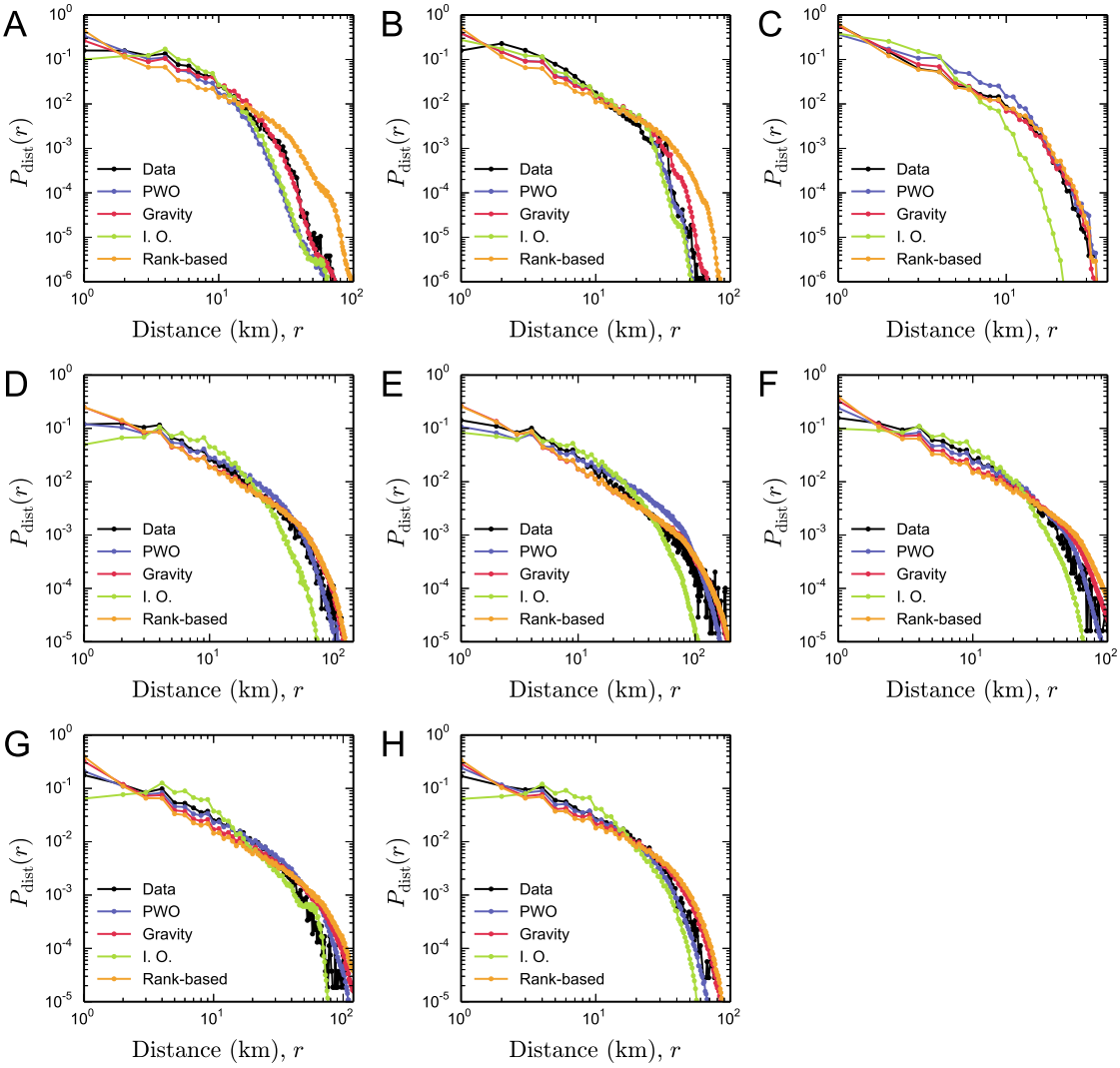


Figure S9: **Comparing the travel distance distributions generated by different models**. (A) Beijing. (B) Shenzhen. (C) Abidjan. (D) Chicago. (E) New York. (F) Seattle. (G) Detroit. (H) Twin Cities.

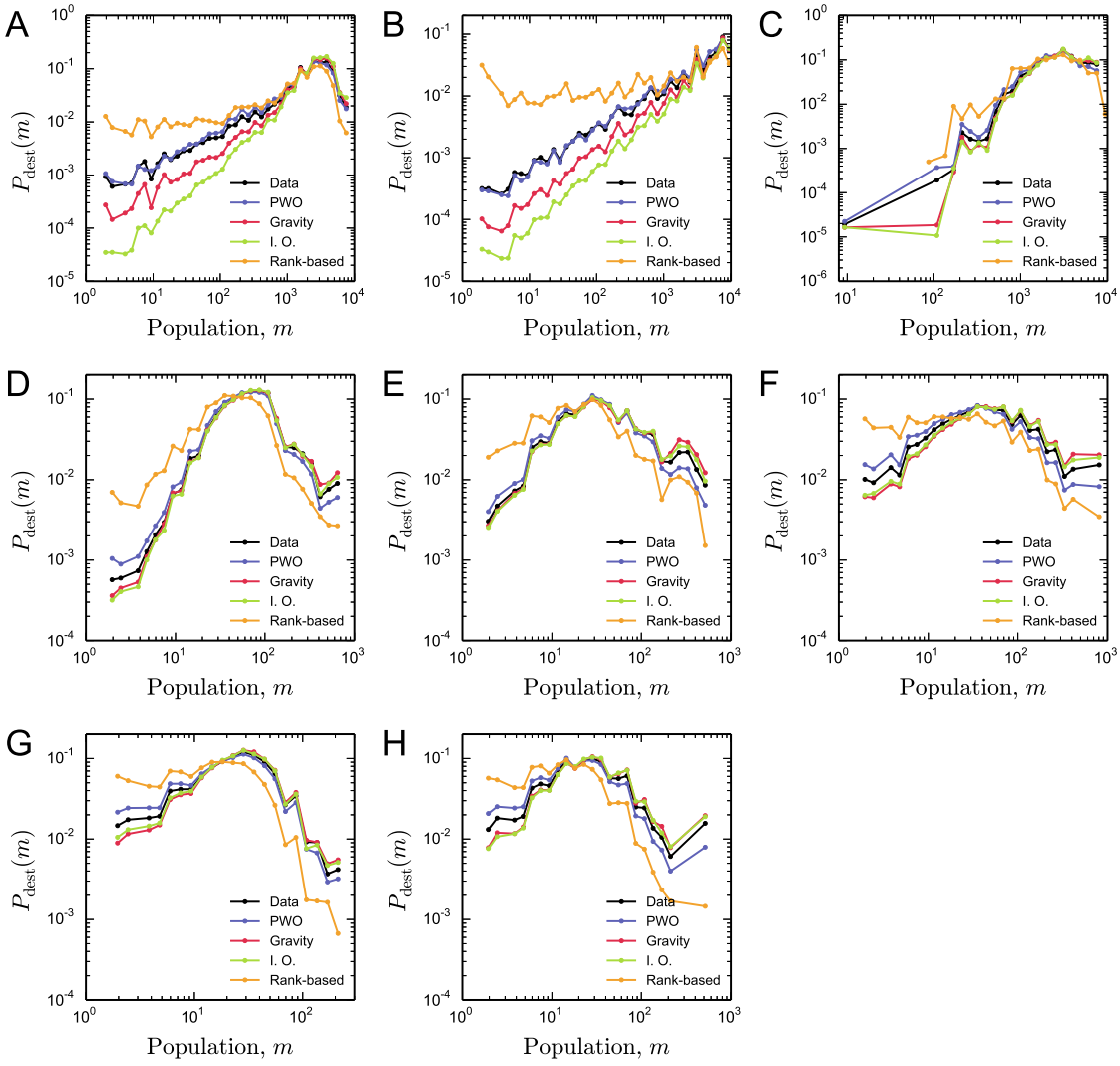


Figure S10: **Comparing the destination travel constraints of different models**. (A) Beijing. (B) Shenzhen. (C) Abidjan. (D) Chicago. (E) New York. (F) Seattle. (G) Detroit. (H) Twin Cities.

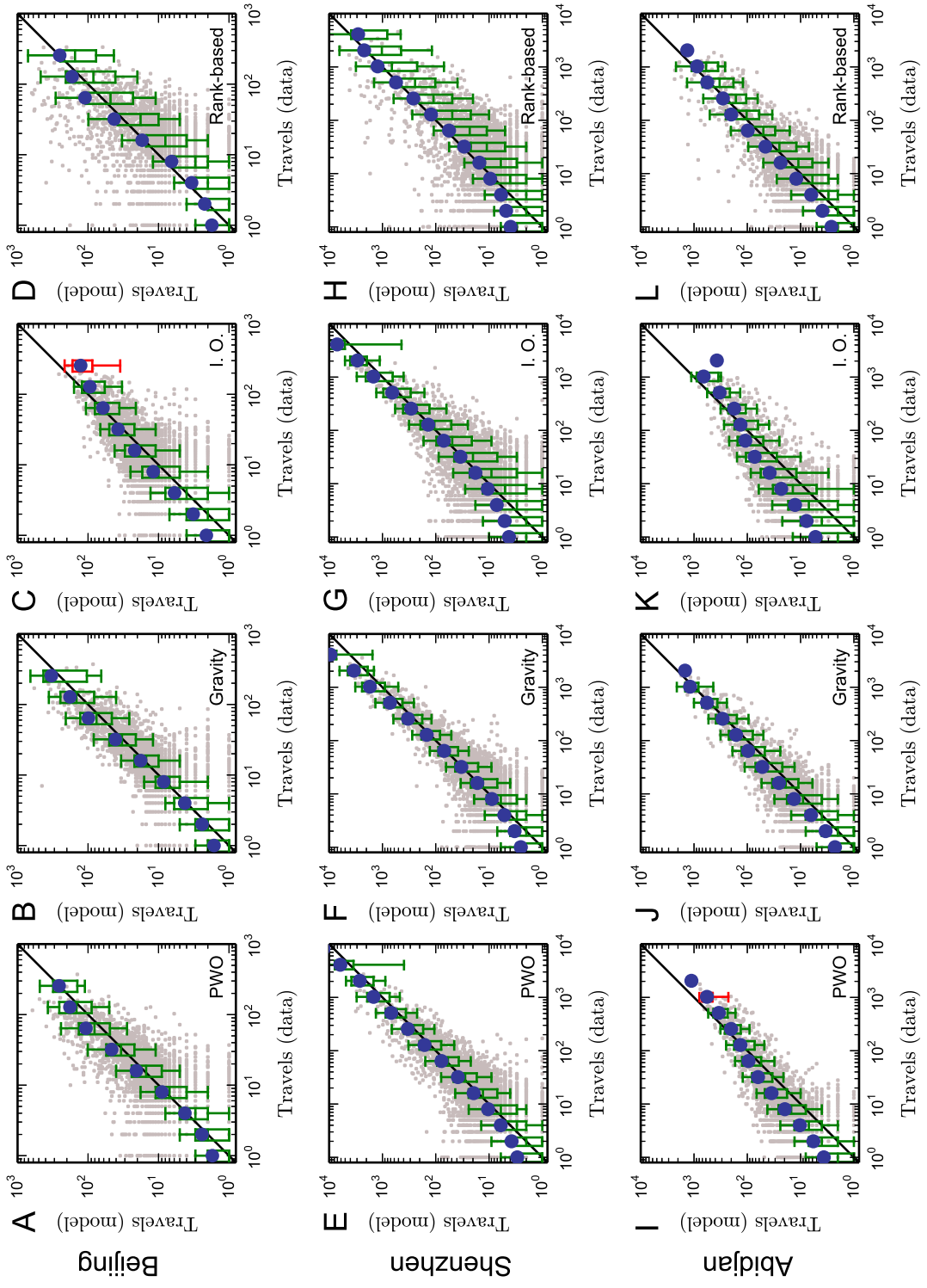


Figure S11: **Comparing the observed fluxes with the predicted fluxes for three Asian and African cities.**

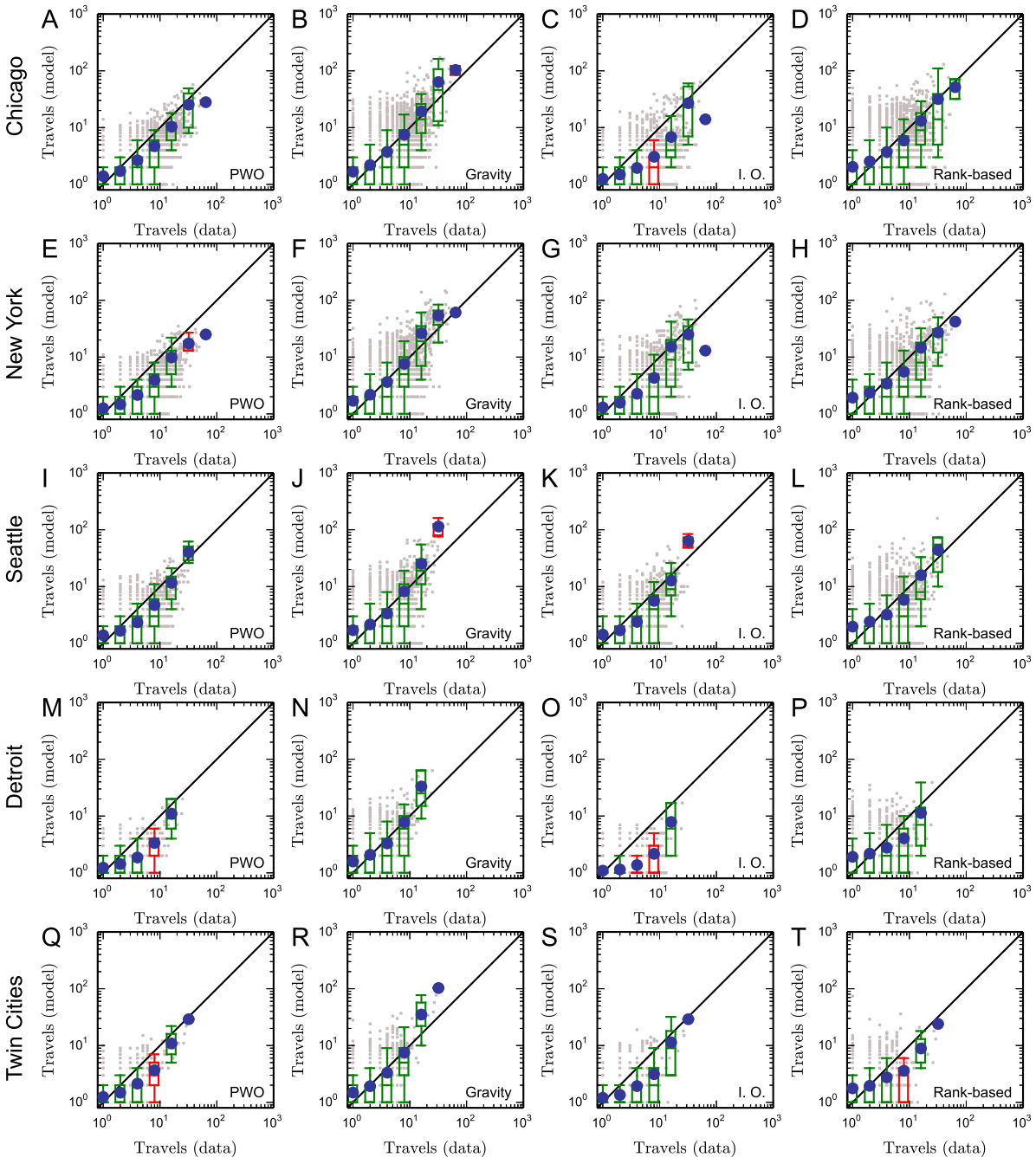


Figure S12: **Comparing the observed fluxes with the predicted fluxes for five U. S. cities**.

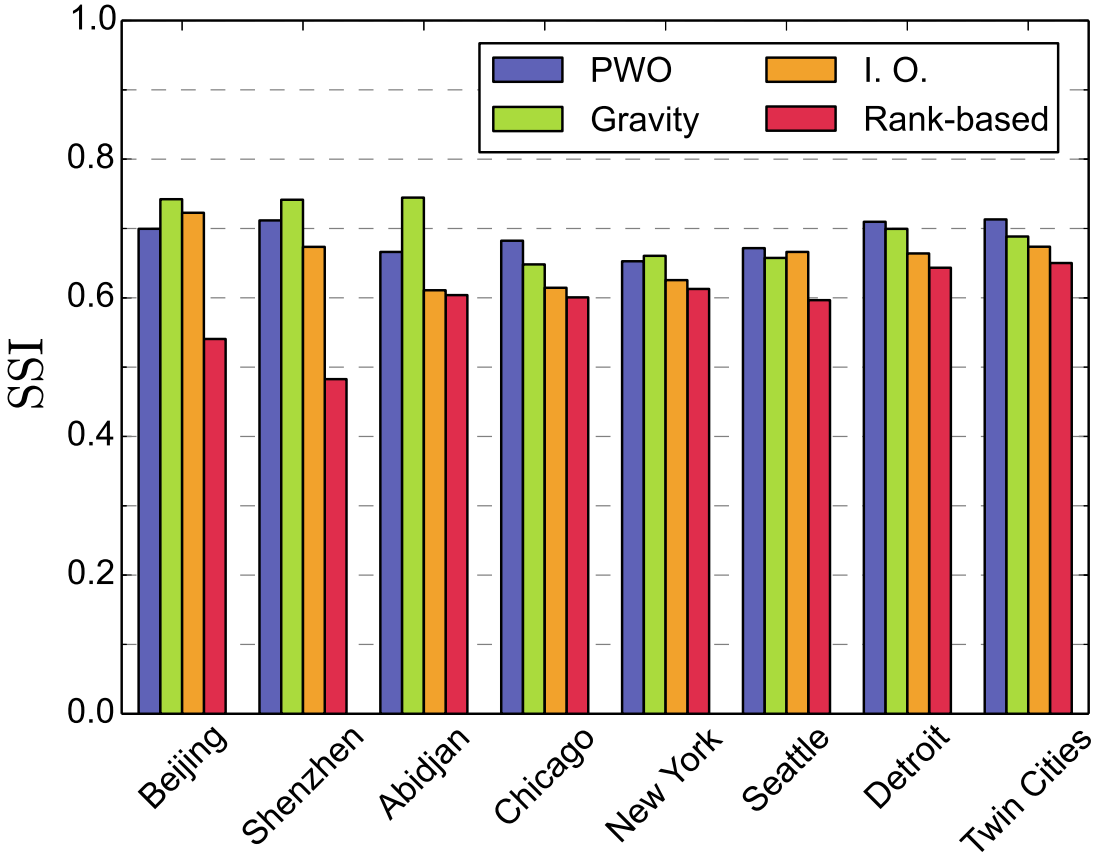


Figure S13: **Comparing the prediction ability of different models based on Sørensen similarity index (SSI)**.

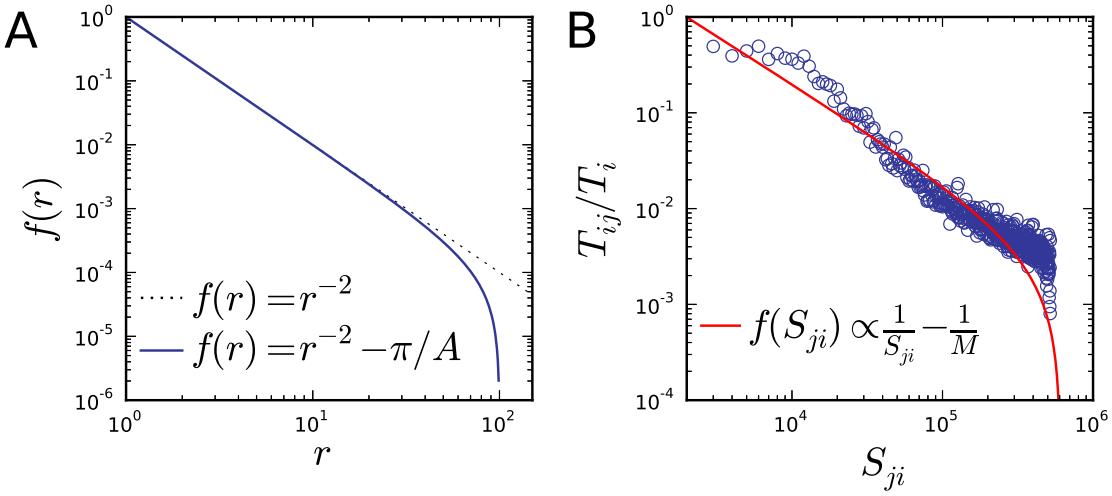


Figure S14: **Distance function and population function**. (A) The distance function $f(r_{ij}) = r_{ij}^{-2} - \frac{\pi}{A}$. $A = \pi 100^2$. (B) The relationship between the population $S_{ji}$ and travel proportion $T_{ij}/T_i$. Blue circles are empirical data and red line is the theoretical function ($M = 607167$).

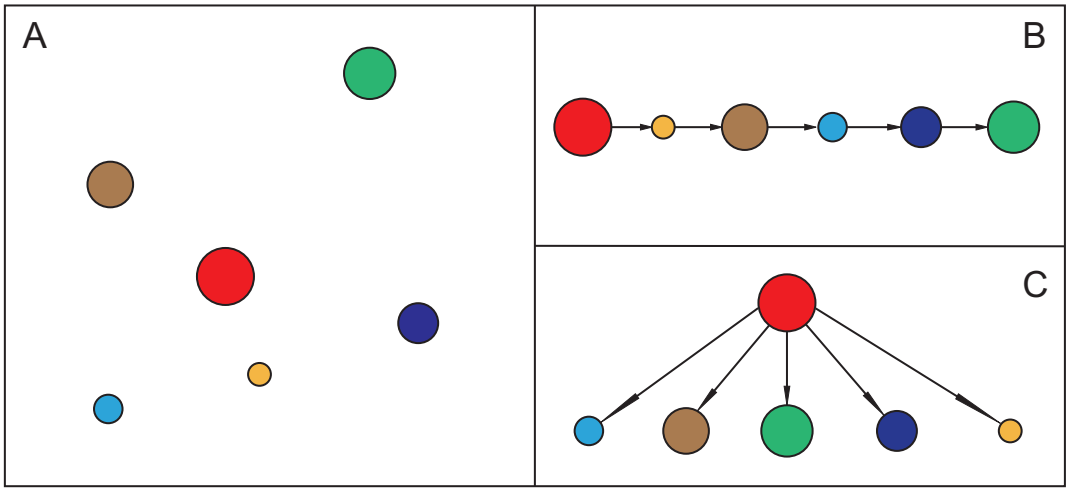


Figure S15: **Schematic description of sequential selection and global selection**. (A) Simplified scenario of destination selection. Central circle (red) is the origin, other circles are selectable destinations. (B) The decision-making process of sequential selection. each traveller ranks all possible destinations in ascending order according to the distance to his/her origin. An individual first decides whether to travel to the first destination in terms of some probability If the individual abandons the first destination, the second one will be considered in terms of the same probability. Subsequently, all possible destinations will be considered step by step until the individual decides to travel to the chosen one. (C) The decision-making process of global selection. a traveller evaluates the attractions of all possible destinations simultaneously and selects a destination to travel with a probability proportional to the destination's attraction.