sustainability

MDPI

Article

Examination on the Influence Area of Transit-Oriented Development: Considering Multimodal Accessibility in New Delhi, India

Sangeetha Ann [1,*], Meilan Jiang [2,*], Ghasak Ibrahim Mothafer [3] and Toshiyuki Yamamoto [3]

1 Department of Civil Engineering, Nagoya University, Nagoya 464-8601, Japan
2 Institutes of Innovation for Future Society, Nagoya University, Nagoya 464-8601, Japan
3 Institute of Materials and Systems for Sustainability, Nagoya University, Nagoya 464-8601, Japan; mdotgh@imass.nagoya-u.ac.jp (G.I.M.); yamamoto@civil.nagoya-u.ac.jp (T.Y.)
* Correspondence: sangeetha.ann@gmail.com (S.A.); jiangml@trans.civil.nagoya-u.ac.jp (M.J.)

Received: 29 March 2019; Accepted: 30 April 2019; Published: 7 May 2019

check for updates

Abstract: As Indian cities adopt the concept of transit-oriented development (TOD), concerns have arisen regarding the applicability of TOD standards formulated in developed countries in the Indian context. This study aims to estimate the TOD influence areas in New Delhi by examining the last mile connectivity patterns of passengers on the Delhi Metro Railway (DMR). Questionnaire surveys conducted on the last mile connectivity reveals use of various access modes for metro stations in India, although current research only considers walking and cycling to be universal forms of access. Therefore, this study focuses on the DMR's multimodal accessibility to investigate the last mile distance of each mode. In order to offset the rounding errors of reported distance, a heaping model and multiple imputation (MI) were employed to improve the accuracy of the reported distance. Afterward, distance decay analysis and receiver operating characteristic (ROC) curves were used to determine the thresholds of last mile distances. The findings show that the influence area differs across travel modes and increases in the order of walking, informal transit, buses, and private transport, respectively.

Keywords: transit-oriented development; multimodal; rounding; multiple imputation; distance decay; receiver operating characteristic

1. Introduction

The catchment or influence area of transit stations is a key factor in transit-oriented development (TOD) planning. There have been many studies on TOD's influence area in developed nations. According to Cervero et al. [1], Bernick and Cervero [2] and Guerra et al. [3], the extent of an influence area is usually based on one's willingness to walk. Researchers have put forward various distances; most range from 400 to 800 m [4–6]. Guerra et al. [3] says that even though 1/2 mile (800 m) has been adopted as the de facto standard for TOD influence areas in the US, "The half-mile transit catchment area, whether radial or network-based is more an artifact of historical precedent than a statistical or analytical construct," raising doubts about the feasibility of adopting this standard. Limited research exists on this regard but show that people are willing to walk or cycle longer than the half mile standard [7–13]. Those research results give different distance values with each other indicating that the transit catchment area may vary among cities, travel modes (both last mile and transit mode), types of area (urban or sub-urban) and trip purposes. Meanwhile, those studies focused only on walking [3,8,10,11] and cycling [12,13] as the access modes.

Recently, TOD is being promoted across India. The Ministry of Urban Development of India gave notification pertaining to the National TOD Policy to promote sustainable urbanization in Indian cities [14]. Based on research from developed countries, India's national TOD policy [14] defines an influence area as being "in the immediate vicinity of the transit station, i.e., within a walking distance, having high density compact development with mixed land use to support all basic needs of the residents." Furthermore, it specifies the influence zone of transit stations as falling within a walking distance of 500–800 m (i.e., a 10–12 min walk). However, it is necessary to establish whether such measurements apply to Indian cities, since their travel patterns and population densities differ significantly from those of developed states. According to Park et al. [7], these distances may be suitable for American cities, but they may not be suited for cities in Europe and Asia. This is corroborated by Sung and Oh [15] whose research showed that TOD planning factors need to be carefully implemented in high density metropolises such as Seoul, South Korea.

Regarding Indian cities, few studies have examined last mile distances and influence areas of transit stations. The existing research illustrates that when people access transit stations, they are willing to walk or cycle longer distances than are mentioned in the national TOD policy [16–20]. Arasan et al. [16] investigated non-motorized transport (NMT) in the city of Tiruchirapalli, and found willingness to walk and cycle to be 1700 and 5200 m, respectively. This study calculated the maximum walking and cycling distances after which people will consider switching to public transit. Therefore, their result shows large distances compared to other Indian case studies. They also indicate that those distances depend on factors such as journey purposes, socio economic characteristics of the travelers, the available travel modes, residential environment, etc. Rastogi and Rao [17] scrutinized the multimodal nature of transit access trips and demonstrated that walking is the most preferred mode for distances under 1250 m. For distances longer than 1250 m, various modes (such as walking, bicycling, auto rickshaw, taxi, bus, and car) were competitive. Since 86% of passengers walk less than 1250 m, Rastogi [18] proposed this figure as the catchment area for suburban rail stations in a case study on Mumbai. In terms of time, Rastogi [19] suggested a 5–10 min walk to transit stations as the criterion area to be developed in line with NMT. Johar et al. [20] administered a survey to bus commuters in New Delhi and looked at their walking distances for trips for different purposes. The mean walking distance ranged from 600 to 700 m for diverse reasons such as work, education, recreation and shopping.

However, these studies only focused on the transit access distance by walking and cycling, although various forms of travel are used for last mile connectivity, including informal modes (such as auto rickshaws, cycle rickshaws, electric rickshaws, Gramin seva, etc.), buses and private modes (such as cars and two wheelers). Informal transit is one of the most popular modes, especially for the middle class in Indian cities. Informal transit has been defined as "public transport services that are provided differently as compared to the typical government-provided bus- and rail-based transport in cities" [21]. They perform a critical role in urban mobility by helping to reduce the demand-supply gap of public transit. Guillen et al. [22] highlighted the importance of informal modes in developing countries and found that commuters tend to use these modes even for a short distance. Besides, buses and private modes are also popular access modes to transit stations in Indian cities. This multimodal accessibility is highly different from developed countries, which have simpler transport modes. Furthermore, Indian cities' high density and income disparities require a myriad of access travel modes.

This study focuses on the multimodal accessibility of transit users in India to examine the appropriate influence area of TOD using a survey data carried out in New Delhi, India. The data set was obtained from a questionnaire survey administered by Delhi Metro Railway (DMR) Corporation (DMRC) to DMR passengers in December 2015. The initial analysis of the data however discovered considerable rounding and heaping in the reported travel distance. Rounding and heaping data may cause substantial errors in the model estimation and lead to biased results.

Studies on the survey data often face such rounding and heaping problems, and have employed various techniques to overcome them [23–25]. Heitjan and Rubin [23] proposed a multiple imputation

(MI) method to correct for heaped age values. According to their study, the imputation task mainly consists of two steps: a modeling task, i.e., creating and estimating a heaping model, to predict missing values from the observed values and model parameters; and an imputation task which formulates the posterior distribution according to the estimation results of the heaping model to draw missing values. Heitjan and Rubin [23] proposed an ordered probit regression heaping model and a simple acceptance rejection procedure for the imputation task to adjust for the missing age data. In their research, they performed a Monte Carlo experiment to verify the estimation and imputation task. Drechsler and Kiesel [24] discuss the potential bias form rounding for income data and indicate that analyzing reported data without any adjustment to account for the rounding will lead to biased results. They corrected the rounding error through imputation and implemented a repeated simulation design to illustrate that valid inferences can be obtained using the imputed data. Recently, Yamamoto et al. [25] applied the heaping model developed by Heitjan and Rubin [23] to solve the heaping issue of reported vehicle kilometers traveled. Since longer vehicle kilometers may raise the probability of higher levels of coarseness, an ordered response model was used to account for the coarseness level of reported vehicle kilometers. Making a comparison with the conventional regression model, they demonstrate that the proposed heaping model is more efficient to investigate the effect of the explanatory variables on vehicle kilometers.

Based on those research works, first, we implemented the imputation task to adjust the rounding errors in our data as described in Section 3. After that, we proceeded with the analysis on the travel distances by using the imputed data to achieve our primary goal to examine the influence area with multimodal accessibility. In the field of the examination on the travel distances, two approaches have arisen. One is the distance decay analysis to catch the relationship between distance and the coverage of travel demand. The other one is the receiver operating characteristic (ROC) analysis to determine the threshold distance for traveling. We implemented both approaches to understand the coverage rate of transit passengers with the travel distance by each mode and the maximum point of distance they would travel. Those works are shown in Section 4.

In the distance decay analysis, Zhao et al. [9] proposed an exponential distance decay function to forecast transit walking accessibility. They concluded that the decay function gave a better prediction compared to traditional buffer analysis and the network ratio method. Their work induced attention regarding to distance decay analysis. Referring to Zhao et al. [9], Kimpel et al. [26] posited a logistic decay function to study catchment areas of bus stops, and inferred that the distance decay function could overcome the overestimation problem associated with buffer analysis. In their model, the decay function is modified to a logistic function in order to reflect a more gradual decline in terms of transit demand at short distances, a steeper decline as the distance approaches one quarter mile, and a more gradual tail. Larsen et al. [27] applied the exponential decay function to a case study of Montréal, Canada and suggested that it could be used to better understand catchment areas and the level of access to services. They examined travel distances by walking and cycling for various purposes such as work, school, shopping, and leisure, while the previous two studies only focused on walking. Since distance decay has received more attention in investigations on travel demand, Halas et al. [28] explored issues pertaining to the shape and parameters of distance decay functions based on daily travel-to-work flows. They put forth a compound power exponential function for the shape of each function to detect different decay tendencies in various cities. This function has been extended to include parameters of population and the number of jobs to demonstrate the high potential of the decay function when applied in various studies. In this study, we applied an exponential function to the decay function to estimate the relationship between travel distance and travel demand coverage as described in Section 4.1.

The ROC analysis is applied broadly in the health and medical fields for diagnostic testing. Recently, researchers have used ROC curves to calculate the threshold distances for people who walk as active travelers against those who use other modes as passive travelers [29–31]. Chillon et al. [29] and Rodríguez-López [30] used ROC curves to study the threshold distances walked to school by children of different age groups. Chillon et al. [31] carried out ROC analysis to find the threshold distances for university students who walk or cycle. Based on their work, we executed the ROC analysis for walking to get the threshold of walking distances as shown in Section 4.2.

This paper is divided into five sections. Section 1 introduces this study. Section 2 gives data description and is divided into two subsections. Section 2.1 talks about the study area and data, and Section 2.2 describes the rounding errors in the reported distance data. The imputation process is employed to correct the reported distance data in Section 3 with two steps: the heaping model to account for the rounding errors in Section 3.1, and the MI process to draw adjusted distance values in Section 3.2. The estimation of the influence areas for different modes using the imputed data is described in Section 4. Section 4.1 describes the distance decay analysis and Section 4.2 describes the ROC analysis. Section 5 presents the conclusions of the study.

2. Data Description

2.1. Study Area and Data

New Delhi has been served by its rapid transit system, the DMR, since 2002; it was built in phases. Currently, its network is 317 km long, with 231 stations. The DMR connects New Delhi to the cities of Bahadurgarh, Faridabad, Ghaziabad, Gurgaon and Noida in the National Capital Region (NCR). It is operated and managed by the DMRC. Figure 1 shows the Delhi Metro route network (Phase I and II) existing at the time of our studied data.

New Delhi is one of the first cities to have adopted TOD in India. For the Master Plan of Delhi (MPD), the Delhi Development Authority (DDA) categorized the city's TOD influence areas into three zones: intense, standard and transition. The intense zone is the most compact, with 300 m around all metro stations, and 800 m around regional interchange stations. The standard zone is 800 m around all metro stations. The transition zone is 2000 m around all metro and regional interchange stations, taking 10 min of cycling into account. In these zones, a mixed land use has been suggested which allows flexibility in the mix of various possible uses, except for polluting and potentially hazardous uses [32]. The MPD-2021 [33] specifies a belt with a width of approximately 500 m on both sides of the Mass Rapid Transit System (MRTS) corridor. These influence areas are planned to promote NMT infrastructure, mixed land use, and to improve first and last mile connectivity to encourage transit usage. These standards need to be verified according to the city's characteristics considering last mile connectivity travel patterns.

The use of land, built environment and route conditions are important factors of accessibility in the planning of TOD. In this study, we did not include the impacts of land use on accessibility or vice versa, since there is lack of those data. Instead, we focused on the last mile distance variations among different access modes analyzing the actual last mile travel patterns reported by the transit users.

The last mile travel data used for this study comes from the questionnaire survey carried out by the DMRC in December 2015, as mentioned in the previous section. This survey contained questions concerning access and egress travel patterns of DMR passengers, such as trip purposes, travel modes, travel distances and time. Passengers' attributes, such as gender, age and income, were also included. This survey was carried out across 131 metro stations covered by DMR during the survey time.

After excluding records for incomplete responses, we obtained a total of 1348 individual samples, with 927 records for access and 858 for egress. The sample size of each station is not adequate enough to carry out analysis on a specific station. The average sample size across all stations was about 8, less than 20 in most stations. Therefore, we conducted a generic analysis across all stations rather than specific station based analysis.

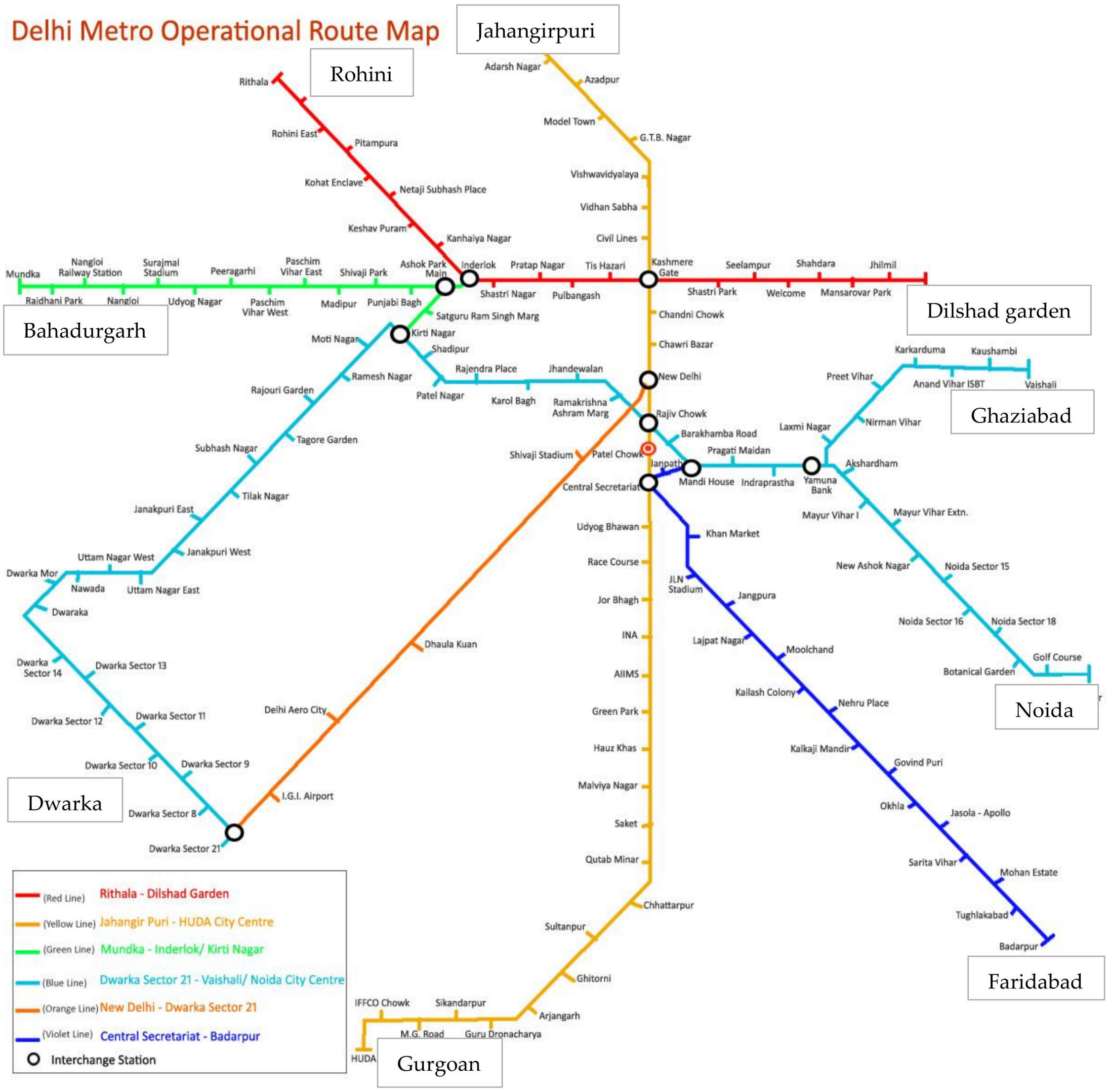
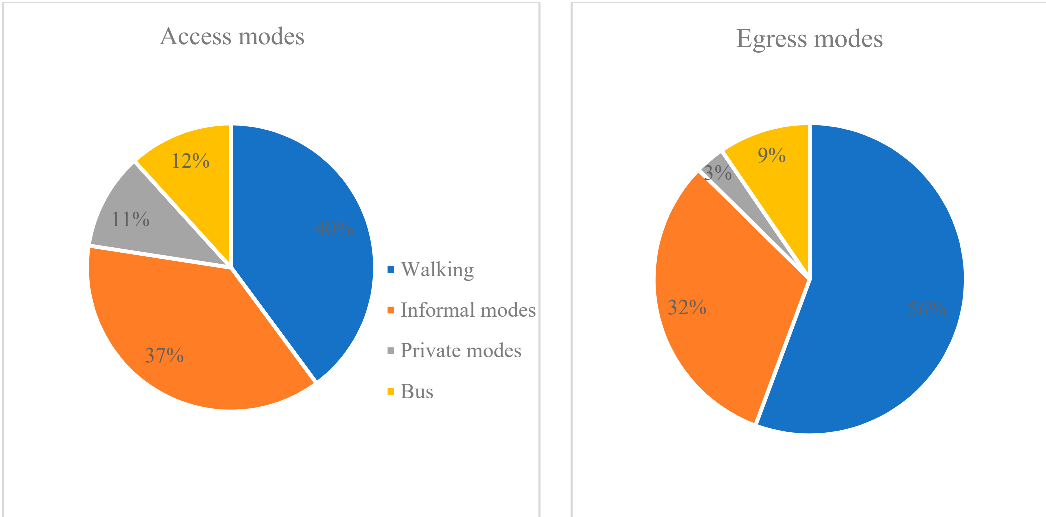


Figure 1. Delhi Metro network route map (Phase I and II). Source: www.ameliabusinessdirectory.com.

Figure 2 displays the shares of various modes for last mile trips. Walking and informal transit are the most commonly used for both access and egress, followed by buses and private transport. Informal transit includes (shared) auto, cycle and electric rickshaws, in addition to a shared minivan service known as Gramin Seva. Private forms include cars and two-wheelers. The share of walking for egress is much greater than for access. The use of private vehicles for egress is much lower compared to access.



(**a**) (**b**)

Figure 2. (**a**) Shares of access modes; (**b**) shares of egress modes.

2.2. *Rounding Problem of Reported Distance Data*

Figure 3 shows the relative frequency distribution and cumulative distribution of access distances by walking. It can be seen that these distances have been heaped at particular points, multiples of 100, 500 and 1000 m. This confirms that the reported distance contains rounding effects. In addition, more observations are reported for shorter (less than 1000 m) versus longer distances. The observations become coarser and clustered around multiples of 500 or 1000 m for longer distances. This implies that the respondents rounded most distances, and the coarseness level of the reported distance increases with the traveling distance. Figure 3 also demonstrates big jumps at multiples of 500 and 1000 m.

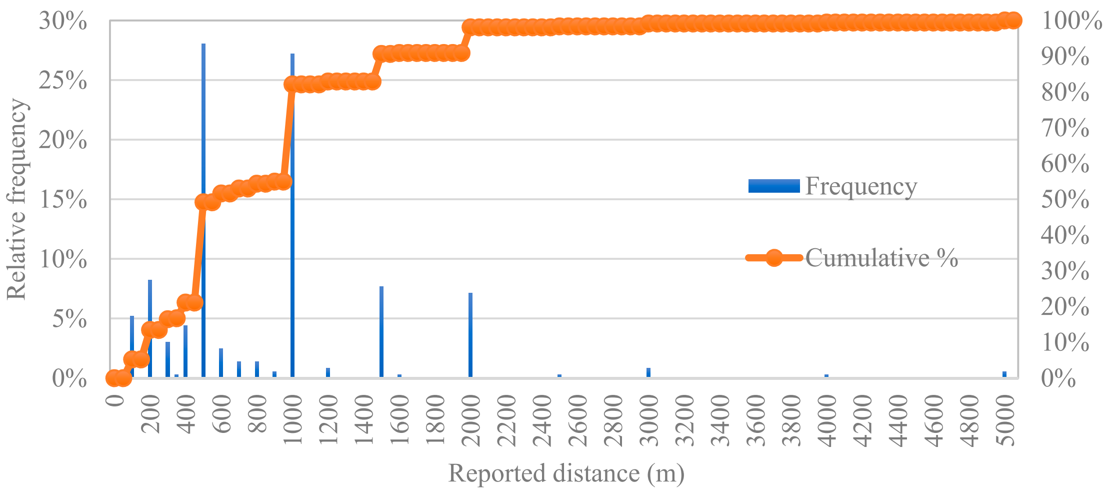


Figure 3. Distribution of reported distances for access by walking.

Table 1 presents sample distributions of reported distance in terms of rounding. The outlined cells indicate the categories with the significant number of rounding. This shows that almost all reported distances are rounded as multiples of 100, 500, 1000 and 5000 m. We observed data rounding at higher numbers (500, 1000 and 5000 m) for private transport and buses, whereas rounding at 100, 500 and 1000 m seems to be popular for walking. This implies that longer distance may raise higher level of coarseness caused by the rounding. Meanwhile, distances traveled by informal modes are rounded at 100, 500, 1000 and 5000 m. We removed bicycling from this study due to its small sample size.

Table 1. Rounding of reported distances for various modes *.

	Cases	Walking	Informal	Private	Bus	Bicycle
Access distances	Multiples of 5000 m excluding multiples of 10,000 m	2	45	19	31	2
	Multiples of 1000 m excluding multiples of 5000 m	129	240	67	69	11
	Multiples of 500 m excluding multiples of 1000 m	131	51	11	7	0
	Multiples of 100 m excluding multiples of 500 m	101	7	0	0	1
	Not multiples of 100 m	1	0	2	0	0
	Total cases	364	343	99	107	14
Egress distances	Multiples of 5000 m excluding multiples of 10,000 m	1	28	8	21	1
	Multiples of 1000 m excluding multiples of 5000 m	137	198	15	54	6
	Multiples of 500 m excluding multiples of 1000 m	183	38	3	5	1
	Multiples of 100 m excluding multiples of 500 m	147	5	0	1	1
	Not multiples of 100 m	4	0	0	0	0
	Total cases	472	269	26	81	9

* The outlined cells indicate the categories with significant number of rounding.

The rounded reported distance may cause errors in our final goal to estimate distances people are willing to travel for access via each mode. In order to account for the rounding errors of reported distance, we applied the imputation methodology to correct the reported distance as described in the following section.

3. Correcting Rounding Errors through Imputation

In order to account for the rounding errors of reported distance, we took two steps: First, we constructed a heaping model to account for the rounding errors of the reported distance data, which formulates a model to predict the coarseness of reported distances. Second, we implemented MI to draw actual distances according to the given distribution of the heaping model.

3.1. A Heaping Model to Account for Rounding Errors

The heaping model to account for rounding errors consists of two functions: a data model function to formulate the distribution of the heaping data and a coarseness function for the rounding behavior. As shown in Figure 3, the reported distance data shows approximately log-normal distribution that is similar to the distribution of vehicle kilometers in Yamamoto et al. [25]. In addition, the coarseness level may increase with longer distance traveled as shown in Table 1. This feature of our data is also similar to the data studied by Yamamoto et al. [25]. Therefore, we applied their heaping model in our analysis.

Our heaping model is built upon Yamamoto et al. [25]. It takes the form of a discrete mixture of an ordered probit model. The model of reported distance is given as:

$$\ln\left(y_i^*\right) = \beta x_i + \varepsilon_i, \tag{1}$$

where y_i^* is the actual distance of individual i, y_i is the reported distance, β is a parameter vector, x_i is a vector of explanatory variables, and ε_i is a random variable that follows a normal distribution. We treated x_i differently for different modes. Based on the cases of rounding shown in Table 1, we assumed walking distances to be rounded to multiples of 100, 500 and 1000 m. We assumed distances traveled informal modes to be rounded to multiples of 100, 500, 1000 and 5000 m. We assumed distances traveled by private transport and bus to be rounded to multiples of 500, 1000 and 5000 m. This means that y_i^* lies in the range [y_i − 50, y_i + 50] if the reported distance is rounded to multiples of 100 m, in the range [y_i − 250, y_i + 250] if multiples of 500 m, in the range [y_i − 500, y_i + 500] if multiples of 1000 m, and in the range [y_i − 2500, y_i + 2500] if multiples of 5000 m.

The levels of coarseness are latent variables and cannot be observed. Therefore, it is not possible to determine whether a distance was rounded to the nearest 100, 500, 1000 or 5000 m. For example, a reported distance of 3000 m could be rounded to multiples of 100, 500 or 1000 m. According to Yamamoto et al. [25], the coarseness of the reported value can be modeled as a latent variable. Since the coarseness level changes with the length of distance, it can be defined as a function of the actual distance. In addition, respondents' socio-economic characteristics may also affect the coarseness of their report. Therefore, the coarseness function is defined as

$$z_i^* = \alpha \ln\left(y_i^*\right) + \gamma X_i + \zeta_i, \tag{2}$$

where, z_i stands for the coarseness of the report, α and γ are parameters, and ζ_i is a normally distributed random variable.

The coarseness of the reported data can be discretized as Equation (3) if there are three multiple cases according to walking, private transport and bus.

$$\begin{aligned} z_i &= 1 \text{if } z_i^* < 0 \\ &= 2 \text{if } 0 \le z_i^* < \theta_1 \\ &= 3 \text{if } \theta_1 \le z_i^*. \end{aligned} \tag{3}$$

On the other hand, it can be extended to Equation (4) if there are four multiple cases, as shown in the informal mode.

$$\begin{aligned} z_i &= 1 \text{if } z_i^* < 0 \\ &= 2 \text{if } 0 \le z_i^* < \theta_1 \\ &= 3 \text{if } \theta_1 \le z_i^* < \theta_2 \\ &= 4 \text{if } \theta_2 \le z_i^*. \end{aligned} \tag{4}$$

Given the coarseness, the reported distance can be represented by the ordered response model with known thresholds. For example, in the case of walking, the reported distance is heaped as multiples of 100 m if $z_i = 1$, multiples of 500 m if $z_i = 2$, and multiples of 1000 m if $z_i = 3$. One of the thresholds is normalized at 0 to identify an intercept term in the latent coarseness process.

Since the coarseness z_i^* depends on $\ln(y_i^*)$, it is assumed to be distributed as a bivariate normal with mean

$$E = \begin{pmatrix} \ln y_i^* \\ z_i^* \end{pmatrix} = \begin{pmatrix} \beta xi \\ \alpha\beta xi + \gamma xi \end{pmatrix}, \tag{5}$$

and covariance matrix V is given by:

$$V = \begin{pmatrix} \ln y_i^* \\ z_i^* \end{pmatrix} = \begin{pmatrix} \sigma_\varepsilon^2 & \alpha\sigma_\varepsilon^2 \\ \alpha\sigma_\varepsilon^2 & \sigma_\zeta^2 + \alpha^2\sigma_\varepsilon^2 \end{pmatrix}, \tag{6}$$

where σ_ε^2 and σ_ζ^2 are variances of ε_i and ζ_i, respectively. In order to identify α, $\sigma_\zeta^2 + \alpha^2\sigma_\varepsilon^2$ is normalized to 1 to fix the scale.

A region $S(y_i)$ of possible values for $(y_i^*,\ z_i^*)$ can be defined, which all map to y_i. It is defined in the case of the informal mode whereby the region $L_i = [y_i - 50,\ y_i + 50) \times (-\infty,\ 0)$ corresponds to heaped multiples of 100 m, $M_i = [y_i - 250,\ y_i + 250) \times [0,\ \theta_1)$ corresponds to heaped multiples of 500 m, $N_i = [y_i - 500,\ y_i + 500) \times [\theta_1,\ \theta_2)$ corresponds to heaped multiples of 1000 m, and $H_i = [y_i - 2500,\ y_i + 2500) \times [\theta_2,\ \infty)$ corresponds to heaped multiples of 5000 m, as shown in Equation (7). The same rule applies to walking, private transport, and buses.

$$\begin{aligned} S(y_i) &= L_i && \text{if } y_i \bmod 100 == 0, \text{ and } y_i \bmod 500 \neq 0 \\ &= L_i \cup M_i && \text{if } y_i \bmod 500 == 0, \text{ and } y_i \bmod 1000 \neq 0 \\ &= L_i \cup M_i \cup N_i && \text{if } y_i \bmod 1000 == 0, \text{ and } y_i \bmod 5000 \neq 0 \\ &= L_i \cup M_i \cup N_i \cup H && \text{if } y_i \bmod 5000 == 0. \end{aligned} \tag{7}$$

The log-likelihood function for the parameters is given as Equation (8) and estimated by the maximum likelihood (ML) method.

$$LL = \sum\nolimits_{i=1}^{n} \ln \int_{S(y_i)} f\left(ln\ y_i^*, z_i^*\right) dy_i^* dz_i^*, \tag{8}$$

where $f\left(\ln\ y_i^*, z_i^*\right)$ is the bivariate normal of E and V.

Table 2 displays the explanatory variables used in the model. The rounding level of reported distance may differ among the respondents. In order to account for differences between them, we studied individual and household characteristics (such as gender, age, household income, individual income and vehicle ownership) as explanatory variables based on the data set. The heterogeneity of the passengers was addressed in the model by including these socio economic variables in the reported distance function as well as in the coarseness function. The vectors x_i in Equation (1) and X_i in Equation (2) include socio economic explanatory variables mentioned in Table 2. Further, the error terms in these functions take into account the unobserved heterogeneity and unpredictability of travel behavior of passengers.

Table 2. Explanatory variables used for the heaping model.

Variable	Definition	Percentage of Value 1 for Access				Percentage of Value 1 for Egress		
		Walking	Informal	Private	Bus	Walking	Informal	Bus
Gender	Dummy: 1 if respondent is male; 0 if otherwise	0.74	0.69	0.80	0.70	0.75	0.67	0.69
Under 30 years old	Dummy: 1 if respondent is younger than 30; 0 if otherwise	0.67	0.72	0.48	0.73	0.66	0.69	0.69
Low household income	Dummy: 1 if monthly household income is less than 30,000 INR; 0 if otherwise	0.29	0.32	0.15	0.29	0.27	0.32	0.23
Low individual income	Dummy: 1 if monthly individual income is less than 30,000 INR; 0 if otherwise	0.74	0.79	0.53	0.76	0.71	0.78	0.81
Vehicle ownership	Dummy: 1 if there is a vehicle in the home; 0 if otherwise	0.75	0.74	0.93	0.81	0.76	0.74	0.96

We estimated this bivariate ordered response probit model using GAUSS [34], a matrix-programming software that provides routines for ML estimation. Table 3 presents the estimation results. We excluded private transport from the analysis of egress due to the small sample size. Only significant variables are left for each mode. In the coarseness function, the estimated values of log-distance coefficient α are positive in all modes with large t-statistics. This verifies that the coarseness level increases with travel distance. The estimate for gender showed a negative value for access by walking, but a positive value for informal access. This implies that men give less coarseness than women in the case of walking, but more coarseness for informal travel. The reason for this difference requires further research. The estimate of low individual income showed a positive value in the case of buses for egress, but was not significant for other cases. The percentage of low individual income among bus users is higher for egress compared to other modes, as shown in Table 2. This suggests that those with low individual incomes may often ride the bus, and that higher frequency of use may lead to more coarseness since users become accustomed to the distance. In the distance function, the estimates of gender and young age are positive and significant in the case of walking and informal transit. This indicates that men or young people travel longer than others in those cases. The estimates for low household income also show positive values in the case of informal and private transport, implying that individuals with low household incomes travel more than others. This is reasonable for private transport since two-wheelers, a popular form of travel among low-income Indians, is included in the private mode. The estimates of vehicle ownership show negative values for walking in terms of both access and egress, suggesting that vehicle owners are less willing to walk. These estimation outcomes reasonably explain the rounding of reported distances, and are expected to properly account for the rounding errors in the analysis of the data set.

Table 3. Estimation results of the heaping model.

Variables	Access				Egress		
	Walking	Informal	Bus	Private	Walking	Informal	Bus
	Estimates (t-stat.)	Estimates (t-stat.)	Estimates (t-stat.)	Estimates (t-stat.)	Estimates (t-stat.)	Estimates (t-stat.)	Estimates (t-stat.)
Coarseness function							
Log-distance (α)	1.081 (15.670)	0.969 (8.505)	0.982 (5.365)	1.101 (9.784)	1.149 (19.241)	1.061 (11.106)	0.809 (5.773)
Constant	−6.346 (−13.133)	−5.908 (−6.207)	−7.098 (−4.614)	−8.379 (−8.514)	−6.836 (−16.425)	−6.400 (−7.383)	−6.128 (−4.875)
Gender	−0.311 (−1.949)	0.346 (1.963)	-	-	-	-	-
Low individual income	-	-	-	-	-	-	0.838 (2.278)
Threshold (θ_1)	1.564 (8.871)	1.499 (8.268)	2.301 (7.463)	2.119 (7.998)	1.712 (9.266)	1.336 (6.434)	2.460 (7.588)
Threshold (θ_2)	-	3.629 (14.441)	-	-	-	3.620 (13.298)	-
Distance function							
Constant	6.301 (46.412)	7.677 (107.859)	8.450 (102.551)	8.126 (85.327)	6.164 (57.366)	7.816 (182.213)	8.370 (72.098)
Gender	0.306 (3.150)	-	-	-	0.214 (2.683)	-	-
Young age	0.262 (3.096)	0.178 (2.267)	-	-	0.212 (3.047)	-	-
Low household income	-	0.173 (2.110)	-	0.635 (2.803)	-	-	-
Vehicle ownership	−0.319 (−3.141)	-	-	-	−0.226 (−2.794)	-	-
Std. deviation (σ_e)	0.736 (24.923)	0.647 (25.284)	0.827 (13.063)	0.828 (12.005)	0.694 (27.797)	0.682 (22.318)	0.949 (10.462)
Sample size	364	343	107	99	472	269	81
Mean log-likelihood	−2.124	−2.435	−2.822	−2.722	−2.173	−2.398	−2.968

3.2. Multiple Imputation to Obtain Exact Values of Reported Distances

The random heaping model described in the previous section helps us to understand the interval into which the true distance y_i^* falls. Given the reported values and estimated parameters of the heaping model, we used MI to draw the missing values of the true distance based on the work of Heitjan and Rubin [23] and Drechsler and Kiesl [24]. Heitjan and Rubin [23] carried out a Monte Carlo experiment to test the effectiveness of their model and the experiment showed that the model was statistically and numerically consistent to retrieve biased data from the rounded data. Later, Drechsler and Kiesl [24] also performed a repeated simulation process by repeating the whole process of sampling, imputing and analyzing data to verify that the imputation method provides unbiased point estimates. According to their verification, we assume the imputation method can draw unbiased imputed values, and does not execute experiment on the evaluation of this methodology. We focused on applying their methodology to generate exact distances for our goal to estimate the influence area of each access mode. The procedure of the imputation task is described as following.

Given the estimated parameters of the heaping model, $\varphi = (\beta, \gamma, \theta, \alpha, \sigma_\varepsilon)$ and the fixed observed data (y_i, x_i), the $(\ln(y_i^*), z_i^*)$ for individuals $i = (1, \ldots, n)$ follows bivariate normal distributions. The value of y_i confines $(\ln(y_i^*), z_i^*)$ to the plausible region defined by Equations (3), (4) and (7) for each mode. We implemented the imputation task using a simple rejection sampling approach:

(1) Draw candidate values for $\left(\ln\left(y_i^*\right)^{imp}, z_i^*\right)$ from a truncated bivariate normal distribution with mean vector (5) and covariance matrix (6) using the estimated parameters φ, where the truncation points are provided by the maximum possible degree of rounding given the reported distance y_i (e.g., for a reported distance value of 500 m with possible degrees of rounding of 100, 500 and 1000 m), $\ln(y_i^*)$ is bounded by ln(250) and ln(750), and z_i^* has to be $(-\infty,\ \theta)$.

(2) Accept the drawn values if they are consistent with the observed rounded distance (i.e., rounding the drawn value y_i^* according to the drawn rounding indicator z_i^* gives the observed distance y_i), and impute $\exp\left(\ln\left(y_i^*\right)^{imp}\right)$ as the exact distant value.

(3) Otherwise, draw again.

Repeating this procedure m times gives m imputed data points that properly reflect the uncertainty of imputation. In this study, we repeated the procedure 1000 times to approximate the true values of distance for each mode. For example, Figures 4 and 5 display the original and imputed data of the informal mode, respectively. Before imputation, all reported distance data were concentrated at multiples of 100, 500, 1000 and 5000 m.

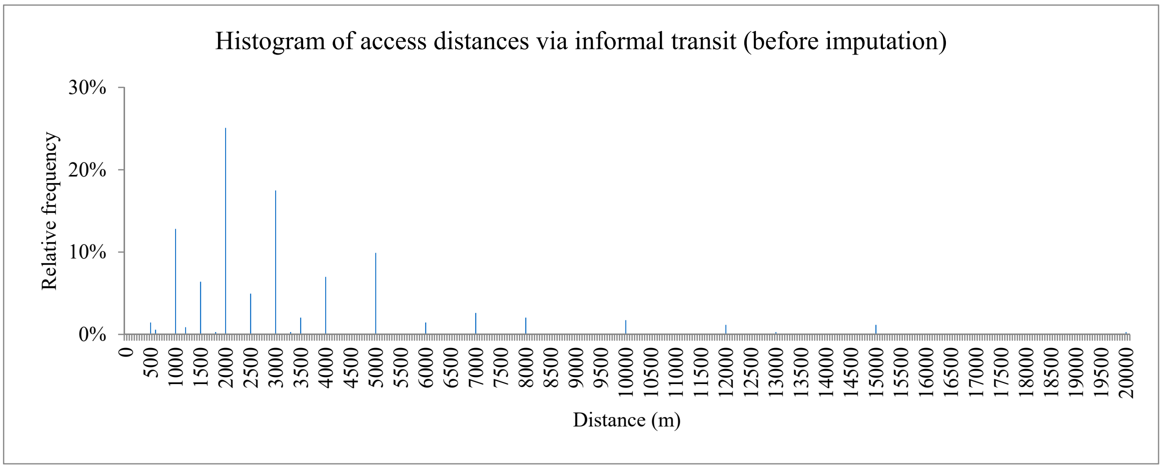


Figure 4. Histogram of access distances via informal transit (before imputation).

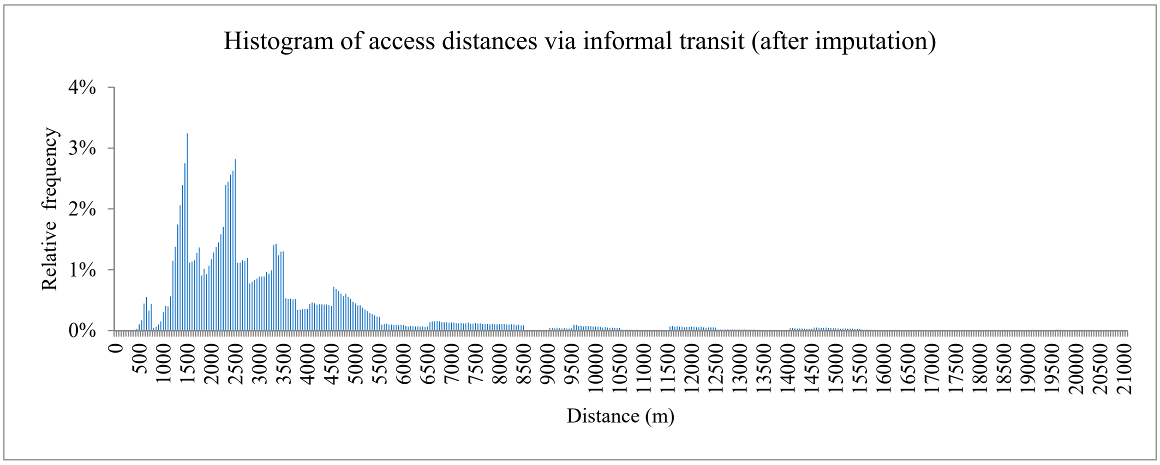


Figure 5. Histogram of access distances via informal transit (after imputation).

This reported distance data also presents relatively high frequencies at 1000, 2000 and 3000 m. After imputation, the rounding effect is accounted for, and imputed distances are distributed around the original report values. Meanwhile, the imputed data properly reflect the relative frequency characteristics of the original data.

4. Estimating Influence Areas of Each Mode

After the imputation, we proceeded with the distance decay analysis for each mode to estimate their influence area. Furthermore, we determined the threshold distances for walking using receiver operating characteristic (ROC) curves.

4.1. Distance Decay Analysis

Distance decay analysis has received more attention in field of investigating distances travelled to access transit stations [9,26–28]. The exponential distance decay function to forecast transit walking accessibility was put forward by Zhao et al. [9].

As shown in Figure 6, after data imputation, it is clear that the coverage of people who walk a given distance (or more) decreases in an exponential shape. Thus, we used the exponential form of the decay function to study travel distances of each mode.

$$y = \exp(-\alpha d), \tag{9}$$

where y is the percentage of passengers traveling a particular distance d (or more), and α is the exponential decay constant to be estimated. According to this decay function, the mean distance can be calculated as $1/\alpha$, and the median value of the distance can be calculated as the 50th percentile of the exponential decay, $\ln(2)/\alpha$.

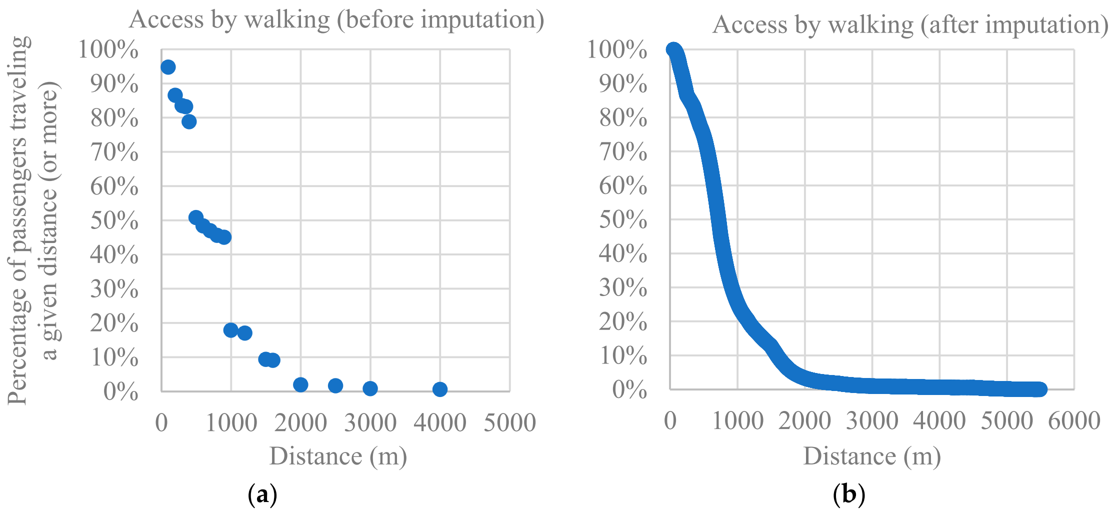


Figure 6. (**a**) Distribution of distances for access by walking before imputation; (**b**) distribution of distances for access by walking after imputation.

We conducted distance decay analysis using the original data set before and after implementing MI. Table 4 summarizes the estimation results of the data set before and after MI for access, while Table 5 covers egress. Since the sample sizes are enlarged after imputation, the modified t-statistic values were used for imputed datasets giving identical sample sizes to before and after MI.

Table 4. Estimation results of decay functions for access.

	Walking		Informal		Bus		Private	
	Before MI	**After MI**	**Before MI**	**After MI**	**Before MI**	**After MI**	**Before MI**	**After MI**
Coefficient α	0.00119	0.00135	0.00030	0.00034	7.70×10^{-6}	0.00016	8.42×10^{-6}	0.00016
(t-stat)	(25.14)	(34.68)	(46.95)	(44.03)	(0.87)	(35.03)	(0.91)	(36.79)
R^2	0.974	0.986	0.991	0.989	0.040	0.985	0.040	0.984
Adjusted R^2	0.915	0.967	0.941	0.983	−0.015	0.982	−0.010	0.982
Sample size	18	18	21	21	19	19	21	21

Table 5. Estimation results of decay functions for egress.

	Walking		Informal		Bus	
	Before MI	**After MI**	**Before MI**	**After MI**	**Before MI**	**After MI**
Coefficient α	0.00136	0.00164	0.00033	0.00038	0.00011	0.00013
(t-stat)	(23.45)	(35.72)	(41.77)	(27.99)	(27.56)	(25.94)
R^2	0.973	0.988	0.991	0.978	0.969	0.964
Adjusted R^2	0.907	0.965	0.928	0.972	0.928	0.962
Sample size	16	16	17	17	25	25

The estimation outcomes of the imputed data show a higher model fit in all modes than data before imputation. In addition, the original data without imputation failed to yield estimations in the case of buses and private transport for access. Moreover, the exponential decay constant α is highly significant across all modes after MI. The estimation results of buses and private transport for access have very close values, implying that they have the same distance decay curve. The exponential decay constant α for walking is significantly higher than that of other modes, showing a rapid fall of coverage percentile with the increase in walking distance.

Figures 7 and 8 show the decay curves estimated from the reported data for access by walking before and after MI respectively. The distribution of imputed data shows a gentle decline at short distances, a shaper decline at a certain distance, and a gradual tail which is consistent with the studies of [11,28]. However, the distribution of original data cannot catch this feature because of the rounding error. The exponential decay function used in our study cannot reflect the gentle decline character of the coverage at short distances properly, while it shows a good fit at long distances. A further improvement in the decay function is required for future study.

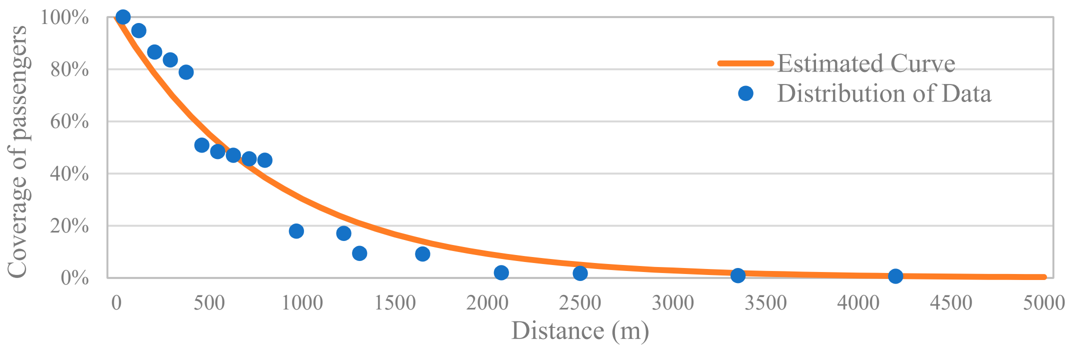


Figure 7. Distance decay curve for access by walking before multiple imputation (MI).

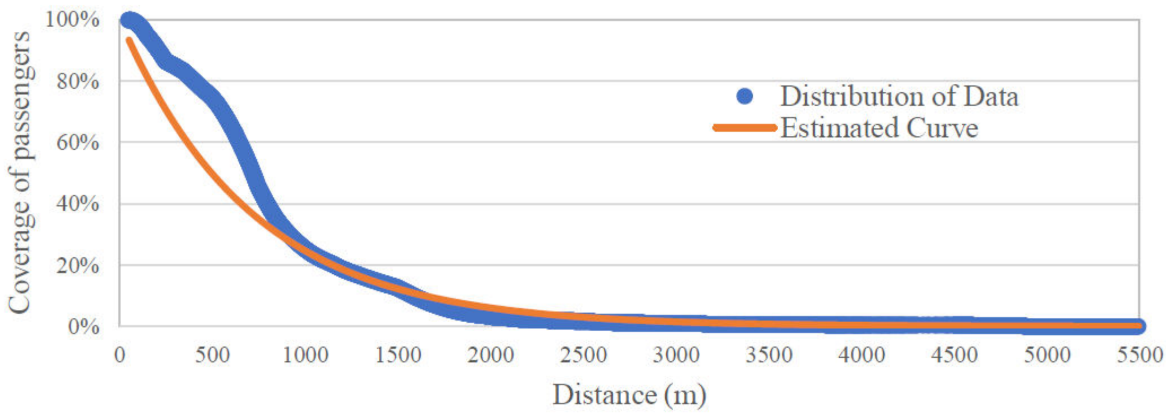


Figure 8. Distance decay curve for access by walking after MI.

Figures 9 and 10 present the decay curves estimated from the imputed data for access and egress, respectively. At the same coverage of passengers, the travel distance increases in the order of walking, informal transit, buses and private transport.

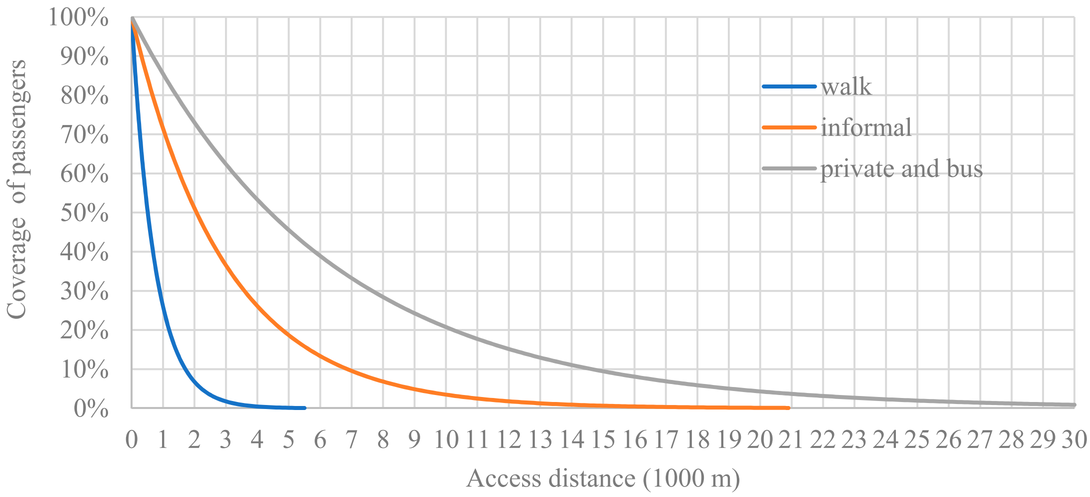


Figure 9. Distance decay curves for all modes of access.

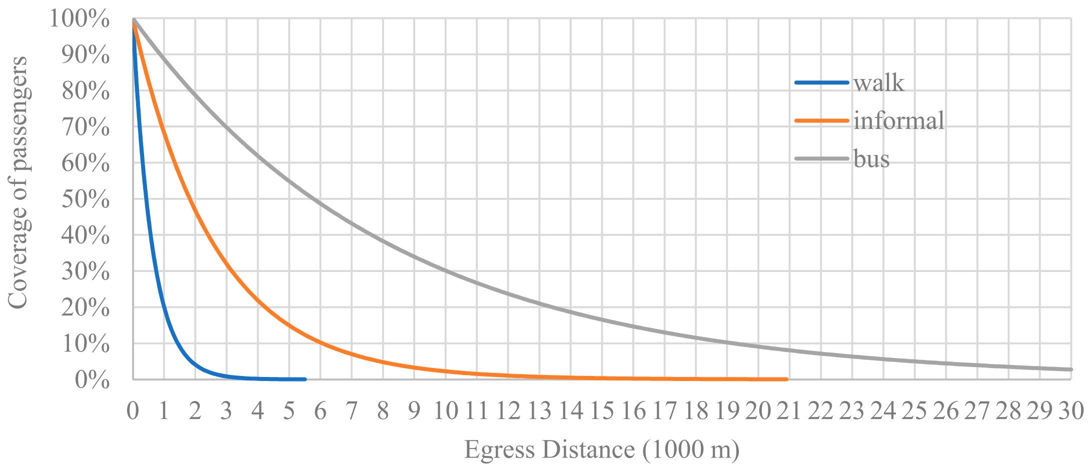


Figure 10. Distance decay curves for all modes of egress.

Tables 6 and 7 summarize travel distances estimated directly from the dataset set as well as from decay functions. The results of before and after MI are also compared. In many cases, the raw data before MI only gives the same distances for different coverage percentiles while the imputed data can catch distance differences among various coverage percentiles. The decay analysis on the raw data can overcome this problem to a certain extent but not for all cases. For example, in the decay analysis of buses and private transport for access, we could not get any useful regression result. On the other hand, the MI can cover the heaping problem properly giving different travel distance values to each coverage percentile through all modes. Furthermore, the decay analysis on the imputed data can detect the different decay tendencies of travel distances in various modes.

Table 6. Summary of travel distances for access (m).

	Walk				Informal				Bus				Private			
	From Dataset		From Decay Function		From Dataset		From Decay Function		From Dataset		From Decay Function		From Dataset		From Decay Function	
	Before MI	After MI	Before MI	After MI	Before MI	After MI	Before MI	After MI	Before MI	After MI	Before MI	After MI	Before MI	After MI	Before MI	After MI
Minimum	100	50	-	-	500	269	-	-	500	253	-	-	500	253	-	-
Maximum	5000	5497	-	-	20,000	20,982	-	-	35,000	37,488	-	-	40,000	42,477	-	-
Mean	800	800	800	700	3300	3300	3300	2900	6600	6600	-	6200	5500	5700	-	6300
Median	600	700	600	500	2500	2500	2300	2000	4000	4400	-	4300	3000	3500	-	4300
70th percentile	1000	900	1000	900	3000	3400	4000	3500	7000	7200	-	7500	5000	5200	-	7500
75th percentile	1000	1000	1200	1000	4000	3900	4600	4100	8000	8100	-	8700	6000	6100	-	8700
80th percentile	1000	1200	1300	1200	5000	4400	5400	4700	10,000	9800	-	10,100	8000	8100	-	10,100
85th percentile	1500	1400	1600	1400	5000	4900	6300	5600	12,000	11,900	-	11,900	12,000	11,700	-	11,900
90th percentile	1500	1600	2000	1700	6000	6000	7700	6800	15,000	13,200	-	14,400	15,000	13,400	-	14,400

Table 7. Summary of travel distances for egress (m).

	Walk				Informal				Bus			
	From Dataset		From Decay Function		From Dataset		From Decay Function		From Dataset		From Decay Function	
	Before MI	After MI	Before MI	After MI	Before MI	After MI	Before MI	After MI	Before MI	After MI	Before MI	After MI
Minimum	50	26	-	-	400	274	-	-	500	255	-	-
Maximum	5000	5483	-	-	15,000	16,000	-	-	35,000	37,498	-	-
Mean	700	700	700	600	3200	3200	3000	2600	7100	7000	9100	7700
Median	500	700	500	400	2000	2400	2100	1800	3500	3500	6300	5300
70th percentile	1000	800	900	700	3200	3500	3600	3200	7000	6700	11,000	9300
75th percentile	1000	800	1000	800	4000	4000	4200	3600	7000	7400	12,600	10,700
80th percentile	1000	900	1200	1000	4000	4400	4900	4200	10,000	8800	14,600	12,400
85th percentile	1000	1000	1400	1200	5000	4800	5800	5000	12,000	12,400	17,200	14,600
90th percentile	1000	1200	1700	1400	6000	6200	7000	6000	20,000	18,900	21,000	17,800

Distances estimated from the imputed dataset were rounded to multiples of 100 m in order to give a reference for TOD policies. The distant curves of private transport and buses are consistent with each other since their exponential decay constants are of the same value.

Comparing access and egress, access distance is slightly longer than egress in the case of walking and informal, while shorter in the case of bus.

The mean distance increases in the order of walking, informal transit, private transport and buses. We thus conclude that faster modes are attributed to longer distances. The estimation result of decay function with the imputed data shows that the mean values of distance for access are 700 m for walking, 2900 m for informal transit, 6300 m for buses and private transport. India's national TOD policy specifies 500–800 m as the extent of an influence area for walking. According to the decay curves, this distance only covers 50%–65% of transit passengers who walk to stations.

The 85th percentile value has been calculated to define catchment areas around transit stations for walking and cycling [9,11,35]. Therefore, we consider the 70th–90th coverage percentile for each mode in order to evaluate the appropriate influence areas. As shown in Table 6, the 85th percentile distances for access are 1400 m for walking, 5600 m for informal transit, 11,900 m for private transport and buses, estimated from the decay functions with imputed data. Informal transit has a shorter distance than private transport and buses, since some informal modes (such as cycle rickshaws) cannot cover distances as long as motorized modes.

As shown in Figure 2, 60% of passengers use modes other than walking for access, and 44% of passengers use modes other than walking for egress around the DMR. It is necessary to take these passengers into account for TOD planning, especially considering India's high density population. According to the estimation results, a multiple range of influence areas can be valid for TOD planning in cities with high density and multimodal accessibility. Depending on the goal of transit coverage, different influence zones can be assigned to various modes. The TOD has been widely accepted as a valid approach for environmentally sustainable city developments. One of the strategies in TODs is to improve the accessibility of transit and to extend the benefits for larger sections of the society. This can be achieved by extending the influence areas properly and thereby benefitting larger areas around the transit station. Figure 11 presents this concept visually. The influence area of a transit station can be considered in terms of multiple ranges of 1200 m for walking, 4700 m for informal, and 10,100 m for private transport and buses according to the 80th percentile decay estimation results on the imputed data. Each range can be a geographical area covered by the exact travel distance of each mode.

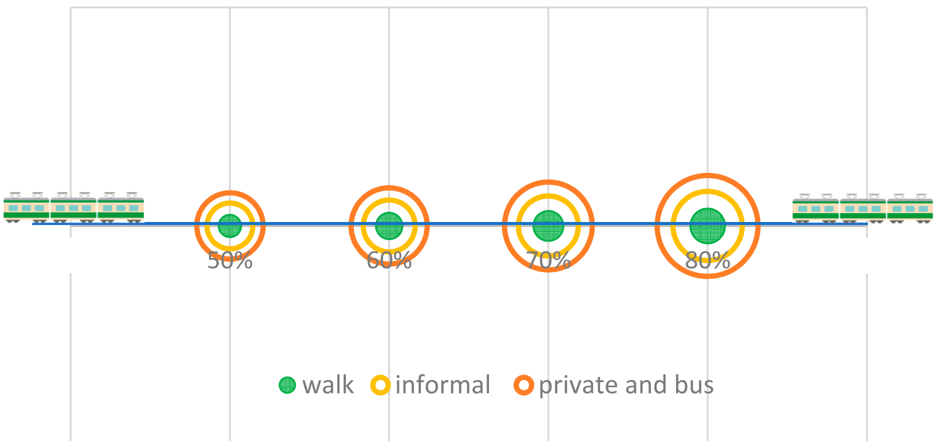


Figure 11. A multiple range of influence areas with different coverages for various modes.

4.2. ROC Analysis

The distance decay analysis provides useful insight into the relationship between the coverage of passengers and travel distance. A criterion for estimating the influence area can be deduced from the decay curves according to the appropriate coverage of passengers in TOD planning. However, decay curves cannot give the threshold for how far people are willing to travel via each mode. In this study, we applied ROC analysis to find the threshold value of distance that passengers are actually willing to walk with imputed data.

The ROC analysis is useful for capturing the trade-off between the true positive rate or sensitivity and the false positive rate (1-specificity) across a series of cut-off points. The threshold values are determined by Youden's index in ROC analysis, highest sensitivity plus specificity, and the closest point on the curve to the (0, 1) corner criterion. According to Gonen [36], the cut-off point can be obtained by realizing the point on the curve with the highest Youden index. Youden's index is the maximum vertical distance from the ROC curve to the line of equality. Meanwhile, the effectiveness of the analysis can be determined by calculating the area under the curve (AUC) (ranging from 0 to 1). The closer the AUC is to 1, the more discriminatory the analysis is [30,31].

In this study, we defined passive walkers as passengers who use informal transit, private transport, and buses instead of walking with a given distance. Figure 12 displays the ROC curves for access and egress for walking before and after MI. The ROC curves estimated with original data have fewer numbers of data points and do not provide a smooth curve like that estimated with imputed data. Due heaping of original data, the Youden's index was the same for multiple data points, the threshold value could not be assigned to a single distance range and the maximum value was therefore considered to be the threshold.

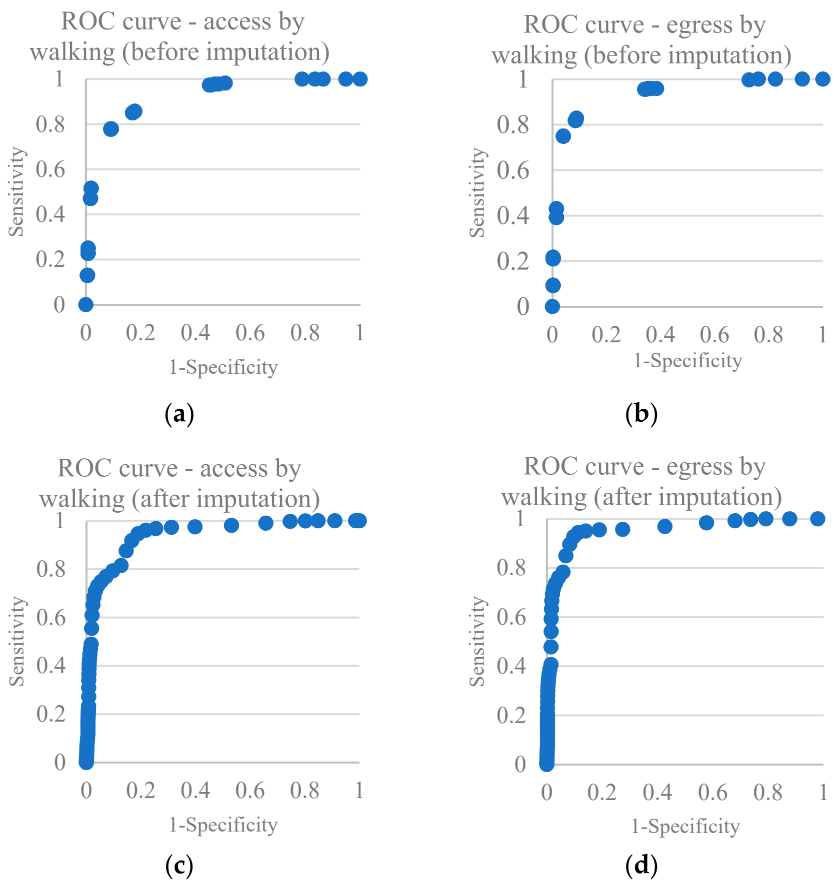


Figure 12. (**a**) Receiver operating characteristic (ROC) curve for access by walking before imputation; (**b**) ROC curve for egress by walking before imputation; (**c**) ROC curve for access by walking after imputation; (**d**) ROC curve for egress by walking after imputation.

Table 8 contains the estimation results. In terms of distinguishing between active and passive participants, we found Youden's index criteria to be 0.757 and 0.834 for access and egress distances after

MI, respectively. The AUC values are close to 1 showing the effectiveness of those results. The threshold distance is around 1200 m for both access and egress. This means that people are willing to walk no more than 1200 m, and most passengers change to another form of travel when the access and egress distances are longer than 1200 m. This distance is close to the 80th coverage distance for access and between 80th and 85th coverage distance for egress, as shown in the decay analysis. The threshold distance calculated using original data for access is around 1700 m and it matched to the 90th coverage distance of the decay analysis. For egress, the threshold value is the same for distance data before and after MI.

Table 8. Estimation results of ROC for walking.

ROC Analysis	Access		Egress	
	Before MI	After MI	Before MI	After MI
Maximum Youden index	−0.688	−0.757	−0.739	−0.831
Threshold (m)	1700	1200	1100	1100
AUC	0.918	0.942	0.932	0.933
Observations N	51	56	51	55

5. Conclusions

In this study, we focused on passengers' multimodal accessibility and last mile distances to DMR stations to examine the appropriate influence area for TOD planning in New Delhi. According to the 2015 survey administered to Delhi metro passengers by the DMRC, walking, informal transit, buses and private transport are considered effective access and egress modes in New Delhi.

The reported distance obtained from such passenger surveys causes considerable rounding and heaping problems, which leads to bias in the estimation of the distance. Therefore, we applied a random heaping model to account for the rounding problem and MI to impute the missing data. We then employed distance decay and ROC analysis to determine the distances travelled by each mode to estimate influence area of each mode.

The estimated distance by decay analysis showed different values for different modes. Travel distance increases in the order of walking, informal transit, buses and private transport. Buses and private transport showed the same decay curve, implying that they cover the same distance from transit stations. In addition, travel distance varies with the coverage percentile of passengers. The result shows that the current TOD strategies, while it specifies 500 m for the influence area, can only cover 50% of the current transit passengers who walk to stations (i.e., 20% of total transit passengers since only 40% of transit passengers walk to stations), and cover very few of the passengers who travel by other modes. This means that almost 80% of current transit passengers would be excluded from the TOD target area. In other words, we need to provide about four times the amount of high density land for a 500 m circle around a transit station in order to cover all transit passengers. This would be unreasonable because Indian cities already have very high density land use, especially around transit stations. Therefore, as it is not feasible to delineate a single influence area, we propose a multiple range of influence areas for TOD planning in Indian cities, considering their multimodal accessibility and high density populations. In addition, the goal of coverage percentage should be carefully specified, taking into account urban population density, population distribution and land use characteristics.

The informal modes are the most preferred mode after walking for the last mile connectivity in New Delhi as shown in the survey data. The distance decay analysis results show that the catchment areas of transit stations with regard to this access mode will be largely extended compared to walking. By including the informal modes, the TOD influence area can be extended to benefit larger areas of the city with sustainable and inclusive development.

Furthermore, we used ROC analysis to check the threshold distance for walking. The threshold value was 1200 m for access and 1110 m for egress, which is consistent with the results of the 80th percentiles of decay analysis. This implies that most passengers are willing to walk longer distances to access transit stations than what is mentioned in the current TOD guidelines for India and the standards adopted by MPD-2021 for New Delhi.

The study approach can be applied in other developing countries especially where there is multimodal last mile connectivity. Many cities in south east Asian countries also have the presence of informal transport modes as well as buses and private vehicles (cars and two-wheelers). These cities can use the empirical findings and the methodology of this study to understand the last mile distances travelled by various modes to transit stations when they explore the concepts of TOD in urban planning.

However, there are some limitations here. We only considered distance because we did not have reliable time data, although time is an important factor in the last mile connectivity. Moreover, this study gives a generic analysis across all stations only considering access and egress distances among different modes. For future work, the effect of land use, built environment and route conditions on the catchment area of TOD can be done focusing on any specific transit station area.

Author Contributions: T.Y., S.A., and M.J. are responsible for conceptualization and methodology. T.Y., M.J., S.A., and G.I.M. worked on software. T.Y., S.A., and M.J. carried out validation and analysis. S.A. and M.J. prepared the original draft. T.Y., S.A., and M.J. reviewed and edited the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors are deeply grateful to the DMRC for all the support and providing the data to carry out this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cervero, R.; Bernick, M.; Jill, G. *Market Opportunities and Barriers to Transit-Based Development in California (Working Paper)*; Transportation Center, The University of California: Berkeley, CA, USA, 1994.
2. Bernick, M.; Cervero, R. *Transit Villages for the 21st Century*; McGraw Hill: New York, NY, USA, 1997.
3. Guerra, E.; Cervero, R.; Tischler, D. The Half-Mile Circle: Does it best represent transit station catchments? *Transp. Res. Board Natl. Acad.* **2012**, *2276*, 101–109. [CrossRef]
4. Calthrope, P. *The Next American Metropolis: Ecology, Community and the American Dream*; Princeton Architectural Press: Princeton, NJ, USA, 1993.
5. Untermann, R.K. *Accommodating the Pedestrian: Adapting Towns and Neighborhoods for Walking and Biking*; Van Nostrand Reginhold: New York, NY, USA, 1984.
6. Dittamar, H.; Ohland, G. *The New Transit Town—Best Practices in Transit Oriented Development*; Island Press: Washington, DC, USA, 2004.
7. Park, S.; Deakin, E.; Jang, K. Can Good Walkability Expand the Size of Transit-Oriented Developments? *J. Transp. Res. Board* **2015**, *2519*, 157–164. [CrossRef]
8. Chia, J.C.S.; Lee, J.B. Variation in the walking time to bus stop by the degree of transit captivity. In Proceedings of the Australasian Transport Research Forum, Sydney, NSW, Australia, 30 September–2 October 2015.
9. Zhao, F.; Chow, L.F.; Li, M.T.; Gan, A.; Ubaka, I. Forecasting transit walk accessibility: Regression model alternative to buffer method. *J. Transp. Res. Board* **2003**, *1835*, 34–41. [CrossRef]
10. Agrawal, A.W.; Schlossberg, M.; Ir, K. How Far, by Which Route and Why? *A Spatial Analysis of Pedestrian Preference. J. Urban Des.* **2008**, *13*, 81–98.
11. El-Geneidy, A.; Grimsrud, M.; Wasfi, R.; Tétreault, J.; Surprenant-Legault, J. New evidence on walking distances to transit stops: Identifying redundancies and gaps using variable service areas. *Transportation* **2014**, *41*, 193–210. [CrossRef]
12. Flamm, B.J.; Rivasplata, C.R. Public transit Catchment areas—The curious case of cycle transit users. *J. Transp. Res. Board* **2014**, *2419*, 101–108. [CrossRef]
13. Martens, K. The bicycle as a feedering mode: Experiences from three European countries. *Transp. Res. Part D* **2004**, *9*, 281–294. [CrossRef]
14. MoUD. *National Transit Oriented Development (TOD) Policy 2015*; Government of India: New Delhi, India, 2015.
15. Sung, H.; Oh, J.T. Transit-oriented development in a high-density city: Identifying its association with transit ridership in Seoul, Korea. *Cities* **2011**, *28*, 70–82. [CrossRef]

16. Arasan, V.T.; Rengaraju, R.; Krishna Rao, K.V. Characteristics of trips by foot and bicycle modes in Indian city. *J. Transp. Eng.* **1994**, *120*, 283–294. [CrossRef]
17. Rastogi, R.; Krishna Rao, K.V. Travel characteristics of commuters accessing transit: A case study. *J. Transp. Eng.* **2003**, *129*, 684–694. [CrossRef]
18. Rastogi, R. Willingness to shift to walking or bicycling to access suburban rail: Case study of Mumbai, India. *J. Urban Plan. Dev.* **2010**, *136*, 3–10. [CrossRef]
19. Rastogi, R. Promotion of non-motorized modes as a sustainable transportation option: Policy and planning issues. *Curr. Sci.* **2011**, *100*, 1340–1348.
20. Johar, A.; Jain, S.S.; Garg, P.K.; Gundaliya, P.J. A study for commuter walk distance from bus stops to different destination along routes in Delhi. *Eur. Transp.* **2015**, *59*, 1–15.
21. Kumar, M.; Singh, S.; Ghate, A.T.; Pal, S.; Ann, S. Informal public transport modes in India: A case study of five city regions. *IATSS Res.* **2016**, *39*, 102–109. [CrossRef]
22. Guillen, M.D.; Ishida, H.; Okamoto, N. Is the use of informal public transport modes in developing countries habitual? An empirical study in Davao City, Philippines. *Transp. Policy* **2013**, *26*, 31–42. [CrossRef]
23. Heitjan, D.F.; Rubin, D.B. Inference from coarse data via multiple imputation with application to age heaping. *J. Am. Stat. Assoc.* **1990**, *85*, 304–314. [CrossRef]
24. Drechsler, J.; Kiesl, H. Beat the heap: An imputation strategy for valid inferences from rounded income data. *J. Surv. Stat. Methodol.* **2016**, *4*, 22–42. [CrossRef]
25. Yamamoto, T.; Madre, J.-L.; de Lapparent, M.; Collet, R. A random heaping model of annual vehicle kilometres travelled considering heterogeneous approximation in reporting. *Transportation* **2018**. [CrossRef]
26. Kimpel, T.; Dueker, K.; El-Geneidy, A. Using GIS to measure the effect of overlapping service areas on passenger boardings at bus stops. *Urban Reg. Inf. Syst. Assoc. J.* **2007**, *19*, 5–11.
27. Larsen, L.; El-Geneidy, A.; Yasmin, F. Beyond the quarter mile: Examining travel distances by walking and cycling, Montréal, Canada. *Can. J. Urban Res. Can. Plan. Policy (Suppl.)* **2010**, *19*, 70–88.
28. Halas, M.; Klapka, P.; Kladivo, P. Distance-decay functions for daily travel-to-work flows. *J. Transp. Geogr.* **2014**, *35*, 107–119. [CrossRef]
29. Chillon, P.; Panter, J.; Corder, K.; Jones, A.P.; Sluijs, E.M.V. A longitudinal study of the distance that young people walk to school. *Health Place* **2015**, *133*, 133–137. [CrossRef]
30. Rodríguez-López, C.; Salas-Fariña, Z.M.; Villa-González, E.; Borges-Cosic, M.; Herrador-Colmenero, M.; Medina-Casaubón, J.; Ortega, F.B.; Chillón, P. The threshold distance associated with walking from home to school. *Health Educ. Behav.* **2017**, *44*, 857–866. [CrossRef]
31. Chillon, P.; Molina-Garcia, J.; Castillo, I.; Queralt, A. What distance do university students walk and bike daily to class in Spain? *J. Transp. Health* **2016**, *3*, 315–320. [CrossRef]
32. UTTIPEC. *Transit Oriented Development: Policy. Norms. Guidelines*; Delhi Development Authority: New Delhi, India, 2012.
33. Delhi Development Authority. *Master Plan for Delhi—2021*; Delhi Development Authority: New Delhi, India, 2007.
34. Aptech. *GAUSS V. 3.2.32*; Aptech Systems: Maple Valley, WA, USA, 1997.
35. Hochmair, H. Assessment of bicycle service areas around transit stations. *Int. J. Sustain. Transp.* **2015**, *9*, 15–29. [CrossRef]
36. Gonen, M. A review of: "ROC curves for continuous data, by W.J. Krzanowski and D.J. Hand". *J. Biopharm. Stat.* **2010**, *20*, 485–487.