



Joint Models for Consideration of Public Transit and Mode Choice for Work Commute

Ganesh Ambi Ramakrishnan¹ · Karthik K. Srinivasan¹ · Surya Pavan Pynda¹

Received: 2 January 2018 / Accepted: 20 March 2020 / Published online: 1 April 2020
© Springer Nature Switzerland AG 2020

Abstract

The mode share of public transit has been decreasing in many Indian cities, while the ownership and usage of personal vehicles continue to grow. It is important to understand the reason behind this decline in transit ridership and frame policies to reverse this trend. Traditional mode choice models assume that all the modes are considered in the choice set while making a decision, which may not be realistic due to the following reasons. Individuals may not consider some modes due to the unavailability, lack of information, or infeasibility, or incompatibility with their travel and activity patterns. By ignoring this consideration process and assuming that all modes are available can lead to mis-specified models and erroneous policy evaluations. Thus, it is important to model both consideration and mode choice outcomes jointly. Accordingly, the main objective of this paper is to develop a joint model for three inter-related choice dimensions: (1) the consideration of bus, (2) the consideration of train, and (3) the primary mode for the home to work commute. These models are estimated using data from a sample of workers in Chennai city and the key explanatory variables influencing them are investigated. A positive correlation between the consideration of bus and train is found. The results also show that consideration of public transport modes influences the mode choice as expected. Furthermore, the coefficient of (log of) consideration probability on choice utility is significantly different from one (an assumption in many implicit availability/perception models). Also, some of the factors affected both the consideration and choice decisions, thus providing evidence that there is a partial (not complete) mediation effect of consideration on choice. Comparison with alternative models shows that neglecting these effects of consideration on choice can lead to biased coefficients, erroneous inferences, and poor model fit. The degree of vehicle availability and accessibility to transit stations exhibit varied effects in the choice of bus and train modes. In addition, their influence differs between the consideration and choice stages. The findings from this study could help decision-makers in developing separate policies for augmenting the consideration of public transit versus those that aim to retain or increase market shares.

Keywords Public transit · Consideration · Mode choice · Bivariate probit · Bus · Train · Joint model · IAPRU

Background and Motivation

India has witnessed an increase in the total number of vehicles by threefold since 1990, while at the same time, there is a drastic decline in the mode share of public transit (< 30% in some cities). These developments have led to an alarming

increase in congestion and emission levels on urban roads. In this context, public transit, which carries a large share of person trips with a very small fleet (< 10,000 vehicles), has a key role in providing sustainable, equitable, and affordable transport in Indian cities. To arrest and reverse the decline in transit shares, there is an urgent need to understand the associated reasons and develop suitable policies. In this context, this paper investigates the key determinants of consideration of public transit (bus and train) and its choice among workers in Chennai city.

Mode choice models play an important role in the demand forecasting process and policy evaluations. Mode choice models are typically modeled using discrete choice models [6–9], some of which have assumed that the choice set is the same for all individuals. This assumption can be

✉ Ganesh Ambi Ramakrishnan
ambi.ganesh@gmail.com

Karthik K. Srinivasan
karthikks@iitm.ac.in

Surya Pavan Pynda
suryapavan.p4@gmail.com

¹ Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India

behaviorally unrealistic as an individual may not consider some modes due to unavailability, lack of information, infeasibility, or incompatibility with activity or travel patterns. Therefore, some studies have suggested that individuals think about only a subset of alternatives from those that are available [1, 2]. The process of narrowing attention from the universal choice set to a subset of feasible options from which the final choice is made is referred to as consideration [2]. By ignoring this consideration process, the assumption that all modes are available can lead to mis-specified models and erroneous policies [3, 4]. For instance, neglecting consideration effects in mode choice models will yield a non-zero probability even for modes that are not considered.

The consideration and mode choice dimensions need to be modeled jointly due to the following reasons: first, the decision to consider different transit alternatives may not be independent of each other. Second, the choice probability must be constrained to be zero for alternatives that are not considered. Third, the set of factors that influence the consideration and choice might be different and some factors might affect both consideration and choice with varying intensity. Fourth, practically, it is important to understand which segments do not consider transit and the associated reasons. This can help in promoting consideration of transit if it is not currently being considered. Among those who consider transit, the focus must be on policies that could increase usage level or frequency (mode shares).

Explicit and implicit consideration frameworks have been developed in the literature for joint modeling of consideration and choice. The main deficiency in the explicit approach is that with an increasing number of alternatives, the number of choice sets grows exponentially which leads to computational intractability, non-convexity, and interpretability issues. In contrast, the implicit availability/perception models circumvent the dimensionality problem, but could suffer from behavioral inconsistency between consideration and choice probabilities. A detailed review of both approaches along with their advantages and disadvantages is presented in the section “[Literature Review](#)”.

Most of the studies on public transit mode choice, however, do not account for both consideration and choice. Also, some of the mode choice studies that include both consideration and choice model are from developed countries and either aggregate bus and train as a single transit mode or focus on only one of these. Therefore, there is a need to investigate factors influencing public transit consideration and choice in developing countries like India while clearly differentiating factors specific to bus and train modes. The role of unique features such as rapid urbanization, increase in income, increase in vehicle ownership, and availability of specific intermediate public transit modes such as autos and shared autos [5] also warrant analysis. Accessibility, transfers, and reliability of transit may also assume significance

in developing countries. Hence, it is important to study the heterogeneity in public transit consideration and choice due to activity characteristics, travel patterns, and last-mile connectivity in Indian cities.

Given these motivations and gaps in the literature, the main objective of this paper is to develop a joint model of three inter-related choice dimensions: (1) the consideration of bus, (2) the consideration of train, and (3) the primary mode for the home to work commute and investigate the key determinants of public transit consideration and choice using empirical data from Chennai city.

A joint discrete choice model system has been proposed for the above dimensions and estimated using data from a sample of workers in Chennai city. The models are used to address the following substantive research issues. With regard to consideration, the issues of interest are: Whether and how the consideration for bus and train are inter-related? What are the unique factors that influence the consideration of bus but not train and vice versa? What are the common factors that affect both? Among these common variables, does the degree of influence vary substantially between bus and train?

With regard to the choice stage, some of the pertinent questions include: What is the role of consideration of public transit modes on primary mode choice for the home to work commute? Does neglecting consideration information in primary mode choice result in loss of efficiency, reduction in goodness-of-fit, or possible bias in estimates? What are the variables that influence only the choice stage and which factors influence both stages? What policies can be specifically used to increase mode share among workers who consider public transit?

This study contributes to the existing literature by proposing a new implicit mode choice model that relaxes the following assumptions:

- (a) independence of consideration across alternatives;
- (b) identical and fixed coefficient of consideration on choice utility;
- (c) full mediation effect of independent variables (that affect consideration) on choice.

Besides, the proposed model is also theoretically consistent as it ensures that the choice probabilities do not exceed the corresponding consideration probabilities.

The rest of the paper is organized as follows. A detailed review of literature related to this study is presented in the section “[Literature Review](#)”. The empirical data are described and exploratory analysis presented in the section “[Data Description and Exploratory Analysis](#)”. The section “[Likelihood Formulation of the Proposed Joint Choice Model](#)” discusses the modeling approach and the likelihood formulation for the proposed joint choice models. The

salient results and inferences from the models are discussed in the section “[Estimation Results and Findings](#)”. In the section “[Implications Due to Policy and Change in Share of Consideration](#)”, the proposed models are applied for illustrative policy analysis under three scenarios. Finally, the major conclusions are summarized in the section “[Summary and Conclusions](#)” and directions for future work are proposed.

Literature Review

This section presents a brief review of the literature related to the objectives of this study, namely, joint models of consideration and choice, empirical findings from consideration models of mode choice, empirical findings from mode choice studies which is followed by the summary of gaps in the literature.

Joint Models of Consideration and Choice

Numerous studies on decision-making in the field of consumer behavior have represented the choice set formation of individuals through a hierarchical or nested process [1, 2]. However, these studies do not deal exclusively with mode choice decisions. These studies have conceptualized the decision-making process of an individual by introducing a universal set that consists of all alternatives. However, not all the alternatives are always available to every individual, leading to a subset called the availability set. The alternatives about which the decision-maker is aware are said to comprise the awareness set [1, 2]. These studies suggest that though an individual may be aware of several alternatives, he/she is likely to choose the best alternative from a reduced subset referred to as the consideration set [1, 2]. Finally, the chosen alternative is selected by evaluating the alternatives included in this set.

Many traditional mode choice models have assumed that the choice sets are fixed and do not vary across individuals [6, 7]. Some of these studies, however, partially account for availability/consideration effects indirectly through explanatory variables such as vehicle ownership, accessibility to public transit, distance to the workplace, etc. Their inclusion in the utility of mode choice may imply that these variables are compensatory in nature. In other words, shortcomings in these variables may be offset by improvements in others. Therefore, such a model can lead to endogeneity, selectivity bias, and possible mis-specification [3, 4]. Also, by modeling the choice in a single stage, the effects of variables affecting consideration and choice stages cannot be segregated, thus limiting the identification of policies for enhancing the consideration of public transit.

To address the above shortcomings arising from disregarding consideration process, Manski [3] proposed an explicit choice set framework for jointly modeling consideration and choice. The first stage involves the generation of choice sets, and the second stage denotes the actual choice conditional on the choice set. This framework assumes that consideration or availability of an alternative is discrete (0 or 1). Furthermore, the choice set for an individual is assumed to be deterministic, but is latent (as it is unobserved to the analyst). The practical difficulty with this framework is the exponential growth in the number of choice sets as the number of alternatives increases leading to increasing computational complexity, non-convexity, and interpretability problems [4, 8].

To address these issues, some studies have modified the explicit Manski framework using random constraints [9], latent choice-sets [10, 11], captivity [12], and attempting to reduce the size of the available choice sets. Other studies utilized explicit unavailability on mode choice to reduce the number of feasible choice sets. These include using distance to transit as a threshold for the availability of transit [13], availability of personal vehicle for the inclusion of carpool, or personal vehicle as an alternative in the choice set [14].

Cascetta and Papola [4] proposed a completely different approach, namely, the implicit availability or perception (IAP) framework to address the dimensionality problem. The main advantage of IAP models is that they obviate the enumeration of choice sets. Instead in these models, a penalty term is added to the utility of an alternative that represents its perceived availability or consideration. Specifically, the penalty term is set as the logarithm of the consideration probability of the corresponding mode. The smaller the consideration probability, the greater is the disutility of choosing the mode. Thus, the penalty is inversely associated with the degree of availability of a mode in the choice set. The penalty term can also be viewed as an instrumental variable that represents an endogenous regressor (consideration) on choice. However, Paleti [8] pointed out that IAP models provide only first-order approximations of Manski’s two-stage framework and can produce biased estimates without higher order corrections.

Another important limitation of the IAP frameworks is that it is not theoretically consistent. Specifically, the choice probability of an alternative may exceed its consideration probability. Most of these studies using the IAP framework have assumed that the sensitivity of consideration on choice is fixed. While recent studies on mode choice have attempted to relax these assumptions [15, 16], they still ignored the interdependence amongst the consideration of various modes. These studies have also specified the mode choice models in such a way that the variables that influence consideration affect choice only

via consideration. This assumption implies only an indirect effect of a consideration variable on choice (i.e., full mediation assumption) and could lead to omitted variable bias if some variables have a partial mediation effect.

Empirical Findings on Consideration of Transit Alternatives

The need to model consideration probabilities is common to both frameworks above. Typically, they are modeled using independent availability models [17] or constrained multinomial logit models [18, 19]. Most of these studies (except [20]) did not account for interdependence amongst consideration of alternatives. Besides, other than through the explicit exclusion of alternatives that are not considered, these studies do not account for any structural or unobserved dependence between consideration and choice.

Due to the lack of explicitly elicited data about which alternatives are actually considered, many studies have treated consideration decision itself to be a latent random variable. With revealed data on consideration, more accurate and precise estimates may be obtained in practice. Along this line, a recent study on characteristics of premium transit services collected revealed data regarding transit awareness, consideration, and usage in major cities in USA [20]. This study proposes a joint choice model incorporating consideration and awareness of bus and train. However, the mode choice model is built by excluding those alternatives that are not considered from the choice set. The transferability of the results and factors to Indian cities remains to be assessed.

Another study, based on a small sample of students from a university, reported that the walking time to bus stops and metro stations have a negative effect on the consideration of public transit. Furthermore, waiting time and travel cost negatively affect the choice of public transit.

In the context of developing countries, very few studies have modeled consideration of public transit. One such study by Kunhikrishnan and Srinivasan [16] used an IAP model to investigate the factors affecting the consideration of bus and train, but did not account for their interdependence. This study showed that workers who return home for lunch are more likely to consider bus but less likely to consider the train. Also, the male workers were less likely to consider bus than female workers.

Empirical Studies on Mode Choice

With regard to worker's mode choice, the role of variables such as travel time, waiting time, and cost have been well established [6, 7]. In-vehicle and out-of-vehicle travel times are widely reported to negatively affect the choice of the bus [6]. Some studies highlight that leg-wise travel time and cost components (including access and egress segments) play a

role in transit choice as they amplify the cumulative burden to commuters. Besides objective level-of-service variables, subjective factors such as comfort also play an important role. For instance, crowding levels affect the value of time and willingness to pay for public transit [21].

Several studies have investigated mode choice decisions in Indian cities (e.g., Agra [22], Chennai [23, 24], Delhi [25, 26], Kolkata [27, 28], Mumbai [25], Rajkot [29], Thiruvananthapuram [30], and Vishakhapatnam [5]). The salient empirical findings from these corroborate the negative influence of journey time (waiting time and in-vehicle time) and delays [5, 26, 28]. Some of the other key determinants of transit mode choice include accessibility to bus stop [28, 31], number of transfers, onboard information [27], crowding [26, 27], seat availability, cleanliness of bus stops [22], and reliability [25]. Also, socio-demographic factors such as gender, income, and vehicle ownership were also influential in the mode choice decision. For example, women were more likely to choose bus than men [7, 32]. Access distance to transit, security for women, and facilities near transit affects the likelihood of reaching transit by walk. Workers who travel long distances were found to prefer buses over informal transit modes (shared auto) [5].

Gaps in the Existing Literature

The above review points to the following gaps in the literature with regard to consideration and choice of public transit, especially in developing countries. Joint models of consideration and choice make strong and restrictive assumptions regarding the independence of consideration, fixed sensitivity of consideration on choice utility, and full mediation of consideration effects. Furthermore, IAP models are also theoretically inconsistent and only approximations to the explicit Manski framework. Due to the lack of explicitly elicited or revealed data about actual consideration, many models treat consideration as a latent variable which can lead to loss of efficiency and goodness-of-fit. Due to these assumptions, the models are unable to differentiate the effect of important explanatory factors (such as accessibility and vehicle availability) on consideration as well as the choice of bus and train. In the developing country context, very few studies have examined the role of contextual decisions (like return home for lunch and need to visit many places), and last-mile connectivity factors on public transit consideration and choice. This study attempts to bridge some of these gaps in the literature.

Data Description and Exploratory Analysis

This study is based on mode choice data collected from a survey of workers in Chennai, India conducted in 2015–2016. Face-to-face interviews were conducted with nearly 804 workers from randomly sampled households belonging to 12 zones in Chennai city. Only workers above 18 years of age who commute to a fixed workplace were surveyed. The survey questionnaire included questions regarding socio-demographic variables (income, vehicle ownership), land-use (facilities near home, access to bus stops and railway stations) and activity patterns (return home for lunch, pick-up or drop-off kids), subjective factors (crowding, comfort, ease of boarding/alighting bus and train, and ease of crossing roads), and contextual variables (travel at odd times, need to visit many places other than workplace, etc.). Data were also collected regarding work travel, public transit usage, and residential location.

The average household size in the sample was 4.00 which is comparable to census values of 4.10 in 2011 [33]. The median age in the sample is 32 years with nearly three-fourths of the sample was below 40 years of age. Nearly 65% of the workers have an undergraduate or higher degree, 10% have a diploma, and 25% have studied up to 12th class. Almost three-quarters of the workers in the sample earn under Rs. 40,000 per month with 33% in the Rs. 10,000—20,000 income bracket. The average vehicle (two-wheeler or car) ownership per household (hh) is 1.49 which shows a considerable growth rate in the previous 8 years (1.26 in 2008 [34]). The vehicle owned is predominantly two-wheeler (1.22 tw/hh) and average car ownership is 0.27 cars/hh.

The personal vehicle is the most commonly used means of travel to work. More than half of the respondents use a two-wheeler (54%) as the primary means of travel for work, whereas only 6% use cars. Nearly 27% of commuters use public transit as the primary mode for work with bus share being 20% and train share is 7%. Intermediate public transport (auto, shared auto, app-based, and cab/taxi services), company buses, and non-motorized modes together contribute the remaining market share (nearly 4% each). Interestingly, the usage of the bus as the primary mode for work

increases from 19% amongst one-worker households to 23% amongst households having more than two workers which is consistent with the reduction in per capita vehicle availability as the number of workers in the household increases.

The respondents were asked which modes were used at least once to go to work in the previous three months. They were also asked for their most frequently used mode for traveling from home to work. This mode is referred to as primary mode hereafter. It is assumed that if bus or train is not used in three months, and then, they are not considered as the primary mode for the work trip. Nearly 52% of the respondents considered bus, whereas 33% considered train. The greater consideration for bus may be due to better accessibility and wider network coverage. The average distance from home to the nearest bus stop was nearly 530 m compared to 3.1 km distance to the nearest railway station. Among those who considered bus, nearly 38% selected it as their primary mode. In contrast, only about 23% of those who considered train actually chose it as the primary mode for work. In contrast, 63% of two-wheeler owners use it as the primary commute mode. Thus, it is clear that a considerable gap exists between consideration share and mode share of public transit.

The data also suggest a strong correlation between the consideration of bus and train (tetrachoric correlation coefficient is 0.5). The statistical significance of this correlation is corroborated by a Chi-square test (Chi-squared observed = 79.92 vs. Chi-squared critical = 3.84, for 1 degree of freedom at 5% significance level). Hence, the consideration of these transit modes needs to be modeled jointly.

The consideration and mode shares are also analyzed based on captivity status. People who either do not own a vehicle or have driving knowledge are referred to as “captive” to non-personal vehicle modes. Individuals with driving knowledge and have a personal vehicle available to them are classified into the choice segment. Respondents with driving knowledge who belong to households with fewer vehicles than workers are categorized as being semi-captive. Table 1 presents the consideration and mode shares among these three segments.

The table shows that only 20% of captive workers did not consider public transit compared to 31% and 46%,

Table 1 Usage and consideration of public transit for work commute amongst captive, semi-captive, and choice segments

	Did not consider public transit (%)	Considered bus or train but did not use either as usual mode of transport (%)	Considered bus or train and used either as usual mode of transport (%)
Captive segment	20	21	59
Semi-captive segment	31	23	46
Choice segment	46	39	15

respectively, from the semi-captive and choice segments. The conditional mode choice shares of public transit given consideration for these three segments are 74% (59% out of 80%) for the captive group, 67% (46 out of 69%) for the semi-captive group, and less than 30% (15% out of 54%) for the choice segment. Thus, the conversion ratio from consideration to usage drops drastically for the choice segment. This suggests that personal vehicle ownership not only reduces the consideration propensity but also the usage of public transit among those who consider it.

The exploratory analysis also suggests that the mode share of public transit for work trips depends on other activity patterns. For instance, while 90% of the sample does not return home for lunch, the transit (bus and train) share in this segment is 30%. In contrast, for the 10% that do return home for lunch, this share drops to 10%. Thus, transit share appears to be influenced by contextual variables that need to be captured by the proposed models. The proposed modeling approach is discussed in the next section.

Likelihood Formulation of the Proposed Joint Choice Model

The objective of this study is to model three inter-related choice dimensions: (1) the consideration of bus, (2) the consideration of train, and (3) the primary mode for the home to work commute.

The consideration of bus and train are binary indicator variables. The primary mode chosen takes the form of a nominal variable and the universal set of alternatives for this choice includes two-wheeler, car, bus, train, auto, shared auto, company bus, walk, and bicycle.

All three choice dimensions are estimated jointly by maximizing the joint likelihood of consideration of bus, consideration of train and primary mode choice conditional on consideration outcomes. The model is joint due to the following linkages between the three choice dimensions: (1) correlation of unobserved terms between the utilities of consideration of bus and train using a bivariate probit model and (2) a structural linkage between consideration and primary mode choice is included by adding the log-transformed values of the probabilities of bus and train consideration as explanatory variables to the mode choice utilities of bus and train, respectively.

A bivariate probit model structure is proposed to capture the correlation between these two decisions. Let C_{bus} and C_{train} be the binary variables indicating the consideration of bus and train, respectively. Let C^*_{bus} and C^*_{train} be the underlying continuous latent propensities of consideration for bus and train, respectively. The relationship between the response and their latent propensities are as follows:

$$C_{bus} = \begin{cases} 1, & \text{If } C^*_{bus} > 0 \\ 0, & \text{Otherwise} \end{cases} \tag{1}$$

$$C_{train} = \begin{cases} 1, & \text{if } C^*_{train} > 0 \\ 0, & \text{Otherwise} \end{cases} \tag{2}$$

$$C^*_{bus} = X_1\beta_1 + \varepsilon_1 \tag{3}$$

$$C^*_{train} = X_2\beta_2 + \varepsilon_2. \tag{4}$$

X_1 and X_2 indicate the set of explanatory factors affecting consideration of bus and train, respectively, and β_1 and β_2 indicate their corresponding coefficients. The error components are assumed to follow a bivariate normal distribution with zero mean and unit variance as shown below:

$$(\varepsilon_1, \varepsilon_2)^T = BVN(0, \Sigma), \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}. \tag{5}$$

As the normal distributions are symmetric about its mean, $-\varepsilon_1$ and $-\varepsilon_2$ also follow normal distributions

$$\therefore \rho_{\varepsilon_1, -\varepsilon_2} = -\rho_{\varepsilon_1, \varepsilon_2}. \tag{6}$$

Let the systematic components of the utilities of consideration of bus and train be V_1 and V_2 as shown below:

$$V_1 = X_1\beta_1, V_2 = X_2\beta_2. \tag{7}$$

The joint probability of consideration of bus and train can be written as follows:

Case 1

$$P(C_{bus} = 0, C_{train} = 0) = P(\varepsilon_1 \leq -V_1, \varepsilon_2 \leq -V_2) = \Phi_2(-V_1, -V_2, \rho) \tag{8}$$

Case 2

$$\begin{aligned} P(C_{bus} = 0, C_{train} = 1) &= P(\varepsilon_1 \leq -V_1, \varepsilon_2 \geq -V_2) \\ &= P(\varepsilon_1 \leq -V_1, -\varepsilon_2 \leq V_2) \\ &= \Phi_2(-V_1, V_2, -\rho) \end{aligned} \tag{9}$$

Case 3

$$\begin{aligned} P(C_{bus} = 1, C_{train} = 0) &= P(\varepsilon_1 \geq -V_1, \varepsilon_2 \leq -V_2) \\ &= P(-\varepsilon_1 \leq V_1, \varepsilon_2 \leq -V_2) \\ &= \Phi_2(V_1, -V_2, -\rho) \end{aligned} \tag{10}$$

Case 4

$$\begin{aligned} P(C_{bus} = 1, C_{train} = 1) &= P(\varepsilon_1 \geq -V_1, \varepsilon_2 \geq -V_2) \\ &= P(-\varepsilon_1 \leq V_1, -\varepsilon_2 \leq V_2) \\ &= \Phi_2(V_1, V_2, \rho). \end{aligned} \tag{11}$$

The above four equations can be rewritten compactly as follows:

$$P(C_{bus}, C_{train}) = \int_{\epsilon_1=-\infty}^{(2C_{bus}-1)X_1\beta_1} \int_{\epsilon_2=-\infty}^{(2C_{train}-1)X_2\beta_2} \phi_2(\epsilon_1, \epsilon_2, \rho) d\epsilon_1 d\epsilon_2 \tag{12}$$

$$P(C_{bus}, C_{train}) = \Phi_2((2C_{bus} - 1)X_1\beta_1, (2C_{train} - 1)X_2\beta_2, (2C_{bus} - 1)(2C_{train} - 1)\rho), \tag{13}$$

where $\phi_2(\cdot)$ represents the bivariate standard normal probability density function and $\Phi_2(\cdot)$ represents the corresponding bivariate standard normal cumulative density function. ρ signifies the correlation between ϵ_1 and ϵ_2 .

In this study, the availability/consideration of all modes other than bus and train is assumed to be deterministic. Specifically, personal vehicle and non-motorized modes are classified either as deterministically available or unavailable in the choice set based on the following criteria.

- (a) Two-wheeler (car) is assumed to be unavailable if the household does not own a two-wheeler (car).
- (b) Similarly, company bus is excluded from the choice set if it is not available for an individual.
- (c) Walk and cycle are excluded if the distance of the workplace exceeds some threshold (9 km and 16.5 km, respectively). The threshold values are taken as the longest commute distance observed in the sample for these modes.
- (d) Shared auto is not considered if it is not used even once in the last 3 months for work.

Note that only the non-consideration of public transportation is treated in a probabilistic manner in this study in view of the focus in this study on consideration and choice of transit as the primary mode.

The third dimension of the proposed joint model is the primary means of travel to work conditional on the probability of the alternatives included in the choice set. The complete choice set for each individual is obtained as a result of consideration of bus and train as well as the availability of other modes based on the criteria discussed above. A multinomial logit model with a log-transformed value of the probability of consideration of public transit modes is used to capture the effect of consideration implicitly on the choice dimension. The conditional probability given the choice set takes the usual logit form for the available alternatives. The estimation of the choice probability conditional on the observed choice set is enforced based on the rules about deterministic availability/unavailability mentioned earlier. The resulting probability ($P_{i|(C_{bus}, C_{train})}$) of choosing

an alternative i amongst k alternatives in the choice set and their corresponding utilities are shown below:

$$U_{i|(C_{bus}, C_{train})} = V_{i|(C_{bus}, C_{train})} + \mu_i \cdot \log(\lambda_i) + \epsilon_{i|(C_{bus}, C_{train})}, \forall i \in k \tag{14}$$

where,

$$\lambda_i = \begin{cases} P(C_{bus} = 1, C_{train} = 0) + P(C_{bus} = 1, C_{train} = 1), & \text{if } i = \text{bus} \\ P(C_{bus} = 0, C_{train} = 1) + P(C_{bus} = 1, C_{train} = 1), & \text{if } i = \text{train} \\ 1, & \text{otherwise} \end{cases} \tag{15}$$

$$P(i|(C_{bus}, C_{train})) = \frac{e^{V_{i|(C_{bus}, C_{train})}}}{\sum_{j=1}^k e^{V_{j|(C_{bus}, C_{train})}}}, \tag{16}$$

where $U_{i|(C_{bus}, C_{train})}$ indicates the utility of an alternative given the choice set and $V_{i|(C_{bus}, C_{train})}$, and $\epsilon_{i|(C_{bus}, C_{train})}$ indicates its corresponding systematic and random components, respectively. $P(C_{bus})$ and $P(C_{train})$ are the marginal probabilities of consideration of bus and train, respectively, and μ_s represent the corresponding coefficients of log of the probability of consideration.

The conditional error terms (random component) given consideration in Eq. 14 are assumed to be independently and identically distributed (IID) as per the Gumbel distribution. The corresponding systematic component is specified as $V_{i|(C_{bus}, C_{train})} = X_{3i}\beta_{3i}$ where X_{3i} indicates the set of explanatory variables affecting the primary mode i and β_{3i} indicate their corresponding coefficients.

Thus, the joint likelihood L and log-likelihood LL for the three choice dimensions for an individual can be written as follows:

$$L(i, C_{bus}, C_{train}) = P(C_{bus}, C_{train}) * P(i|C_{bus}, C_{train}) \tag{17}$$

$$L = \prod_i L(i, C_{bus}, C_{train})^{\delta_{i|(C_{bus}, C_{train})}} \tag{18}$$

$$LL = \sum_i \delta_{i|(C_{bus}, C_{train})} * \{ \ln(P(C_{bus}, C_{train})) + \ln(P(i|C_{bus}, C_{train})) \}, \tag{19}$$

where $\delta_{i|(C_{bus}, C_{train})}$ represent indicator variables for the choice of the i th alternative as the primary mode for work. All the parameters are estimated simultaneously by maximizing the joint log-likelihood expressed in Eq. 19. The user-defined LL function written in R programming language is used for estimation.

Not accounting for consideration in the mode choice model could lead to omitted variable bias due to common factors influencing the consideration and choice decisions. On the other hand, including the indicators for consideration directly in the mode choice model could lead to endogeneity due to correlation between the unobserved

terms and the explanatory factors in the mode choice utilities. Hence, the log-transformed variables are chosen as instrumental variables to capture the role of consideration of bus and train in the mode choice utilities while trying to minimize potential endogeneity issues.

Furthermore, the presence of heteroskedasticity and correlation across primary mode choice alternatives was also tested separately using mixed logit as well multinomial probit structure. No significant evidence of heteroskedasticity or correlation of unobserved terms of bus and train mode choice utilities was observed for two specifications of interest (with and without the instrumental variables of consideration as explanatory variables). Therefore, a multinomial logit model has been used due to its simplicity, computational tractability, and parsimony.

Thus, all three dimensions are jointly modeled as noted above. However, since the two response dimensions

(consideration and primary mode chosen) are different from each other, they have been presented in separate subsections for ease of exposition and understanding.

Estimation Results and Findings

With regard to consideration, two models M1a (independent consideration) and M1b (correlated consideration) are estimated and compared in Table 3. With regard to mode choice, three models are built and evaluated: M2 (joint model assuming that both transit modes are always considered), M3 (joint model with an implicit choice set assuming full mediation), and M4 (joint model with an implicit choice set assuming partial mediation), as shown in Table 4. M2 assumes that consideration has no role on choice, whereas in M3, the factors that affect consideration only indirectly

Table 2 Description of various specifications estimated for primary mode choice dimension

Model name	Model scenario	Specification of the primary mode choice
M2	No mediation scenario (base specification)	Neither log-transformed probabilities of consideration nor the vector of variables affecting consideration influence primary mode choice
M3	Full mediation scenario	Only the log-transformed probability of consideration affects primary mode choice
M4	Full and partial mediation scenario	Both the log-transformed probabilities of consideration as well as the vector of variables affecting consideration influence mode choice

Table 3 Comparison of results of bus and train consideration component of the joint model

	(M1a) Independent binary probit models				(M1b) Bivariate probit model			
	Bus		Train		Bus		Train	
	coef	tstat	coef	tstat	coef	tstat	coef	tstat
Constant (mode considered = 1)	<i>-0.04</i>	<i>-0.48</i>	-0.69	-9.26 ^a	<i>-0.07</i>	<i>-0.92</i>	-0.69	-10.06 ^a
Distance to work < 2 km	<i>-0.78</i>	<i>-4.82^a</i>	-0.66	-3.53 ^a	<i>-0.75</i>	<i>-4.64^a</i>	-0.70	-4.00 ^a
Direct bus service is available	0.24	2.54 ^a	-	-	0.32	3.53 ^a	-	-
Information regarding bus timings is available	0.47	2.82 ^a	-	-	0.35	2.27 ^a	-	-
Return home for lunch	<i>-0.34</i>	<i>-2.25^a</i>	-0.40	-2.42 ^a	<i>-0.36</i>	<i>-2.71^a</i>	-0.41	-2.79 ^a
Worker belongs to captive segment	0.62	4.77 ^a	-	-	0.56	4.78 ^a	-	-
Worker belongs to semi-captive segment	0.34	2.52 ^a	0.29	2.21 ^a	0.28	2.12 ^a	0.25	1.90 ^a
Income > Rs.40,000 per month	<i>-0.72</i>	<i>-5.72^a</i>	-	-	<i>-0.66</i>	<i>-5.58^a</i>	-	-
Access to railway station < 1.2 km from home	-	-	0.18	1.76 ^a	-	-	0.21	2.45 ^a
Egress to railway station 1 km from office	-	-	0.31	2.77 ^a	-	-	0.40	4.56 ^a
Walking or crossing roads is easy	-	-	0.41	4.20 ^a	-	-	0.37	4.55 ^a
Correlation between bus and train consideration	-	-	-	-	-	-	0.53	9.63 ^a
No. of observations	803				803			
Initial log likelihood	-1113.19				-1113.19			
Converged log likelihood	-976.46				-935.41			
No. of parameters	15				16			

Italicized coefficients are insignificant at 10% level

^aSignificant at 10% level (one- or two-tail)

Table 4 Comparison of mode choice component of the joint model with and without the information of choice set

Variable	Alternative	Model M2		Model M3		Model M4	
		coef	tstat	coef	tstat	coef	tstat
Constant	tw	Base		Base		Base	
	Car	6.82	4.99 ^a	6.14	4.53 ^a	6.21	4.56 ^a
	Bus	-1.48	-4.59 ^a	0.84	1.69 ^a	-0.70	-1.47 ^a
	Train	-2.91	-10.08 ^a	<i>-0.41</i>	<i>-0.68</i>	-0.93	-1.70 ^a
	Auto	2.66	0.65	4.74	1.12	4.78	1.11
	Shau	-1.44	-1.15	-0.71	-0.55	-0.71	-0.54
	Cbus	-3.21	-7.25 ^a	-3.01	-6.38 ^a	-3.04	-6.42 ^a
	Walk	0.15	0.36	-0.51	-1.22	-0.41	-0.96
	Cycle	-2.00	-4.21 ^a	-2.03	-4.42 ^a	-2.03	-4.38 ^a
Travel time	tw, car	-0.04	-6.88 ^a	-0.04	-5.98 ^a	-0.04	-6.03 ^a
	Bus	-0.02	-4.13 ^a	-0.02	-3.58 ^a	-0.02	-3.70 ^a
	Train	-0.01	-3.10 ^a	-0.01	-2.39 ^a	-0.01	-2.49 ^a
	Auto	-0.12	-2.83 ^a	-0.13	-2.93 ^a	-0.13	-2.96 ^a
	Shau	-0.03	-2.64 ^a	-0.03	-2.44 ^a	-0.03	-2.50 ^a
	Walk, cycle	-0.09	-5.81 ^a	-0.08	-5.31 ^a	-0.08	-5.27 ^a
Travel cost	Car	-1.02	-4.55 ^a	-0.92	-4.20 ^a	-0.93	-4.21 ^a
	Bus, train, cbus	-0.12	-1.73 ^a	-0.12	-1.39 ^a	-0.12	-1.44 ^a
	Auto, shau	-0.32	-1.31 ^a	-0.46	-1.77 ^a	-0.46	-1.76 ^a
Income < 20,000	Car	-1.66	-2.84 ^a	-1.59	-2.73 ^a	-1.60	-2.73 ^a
	Bus, cycle	0.39	1.96 ^a	-	-	-	-
	Auto	-1.70	-2.08 ^a	-1.71	-2.10 ^a	-1.72	-2.10 ^a
No. of two-wheeler	Car	-0.70	-2.78 ^a	-0.65	-2.52 ^a	-0.67	-2.59 ^a
	Bus, cycle	-0.52	-3.67 ^a	-0.21	-1.42 ^a	-0.32	-2.05 ^a
Female	Bus, train	0.84	3.58 ^a	0.78	-3.07 ^a	0.83	3.24 ^a
	Auto, shau, cbus	1.82	5.37 ^a	1.85	-5.38 ^a	1.87	5.41 ^a
Distance to work < 2 km	Bus	-	-	-	-	-1.01	-1.41 ^a
Direct bus service is available	Bus	-	-	-	-	0.40	1.42 ^a
Information regarding bus timings is available	Bus	-	-	-	-	0.52	1.37 ^a
Worker belongs to semi-captive segment	Bus	-	-	-	-	1.22	4.08 ^a
Income > Rs. 40,000 per month	Bus	-	-	-	-	-1.31	-1.97 ^a
Worker belongs to semi-captive segment	Train	-	-	-	-	0.75	1.74 ^a
Log-transformed (consideration probability)	Bus	-	-	4.29	5.89 ^a	2.05	4.49 ^a
	Train	-	-	2.41	4.28 ^a	2.05	4.49 ^a
Summary statistics							
Number of observations		803		803		803	
Joint converged Log-likelihood of consideration and choice		-1700.16		-1646.32		-1638.33	
No. of parameters in the mode choice model		24		25		30	

Tw two-wheeler, shau shared auto, cbus company bus

^aSignificant at 10% level (one- or two-tail), italicized coefficients are insignificant at 10% level

influence choice through the penalty terms [$\log P(C_{bus})$ and $\log P(C_{train})$]. M4, on the other hand, captures both the direct and indirect effect of factors influencing consideration on choice. Table 2 shows the differences in the three models associated with the primary mode choice dimension.

The following features of the proposed model (M1b together with M4) enable it to relax the limitations of existing IAP models. First, the use of bivariate probit (M1b)

captures the correlation of consideration for bus and train. Second, the specification in Eq. 14 permits the coefficient of log of consideration probability to be different from 1 and is not necessarily identical across alternatives. Third, the specification in M4 permits some factors to influence choice directly in addition to their influence via the consideration probabilities, thus relaxing the full mediation assumption. Finally, Eq. 17 ensures that the implicit choice probability

is conditional on consideration, and the unconditional probability cannot exceed the consideration probability for any alternative.

The results obtained from the joint model have been tabulated separately into two tables (Tables 3, 4) for ease of presentation of the findings from consideration and primary mode choice. The proposed model accounts for correlation of unobserved terms between consideration of bus and train through the correlation coefficient. On the other hand, the significance of the log-transformed probabilities of consideration captures the structural correlation between consideration and primary mode choice.

However, MNP and mixed logit models were developed to test possible heteroskedasticity or correlation between bus and train utilities, but turned out to be insignificant. Hence, the “rho” (i.e., correlation of error terms) has been reported only in Table 3, while the coefficient for the log-transformed probability is reported in Table 4.

Correlation in Consideration of Bus and Train

The results from Table 3 show that the correlation between consideration of bus and train is significant (0.53 with t-stat of 9.63). Capturing this correlation results in improving the log-likelihood of consideration decisions from -976.46 for the independent model (M1a) to -935.41 for the correlated model (M1b). This is consistent with exploratory analysis results, indicating that there are common unobservable factors that influence the consideration of both bus and train, possibly because they are both scheduled modes.

A comparison of magnitudes, signs, and significance of explanatory variables across the models showed that only a few coefficients were different across the models. The effect of distance to work on train consideration, and that of direct bus service on bus consideration were underestimated by

a little more than 10% in the independent model, whereas the effect of accurate information on bus routes and timings were overestimated by a similar margin.

However, the distribution of estimated consideration probabilities (in Fig. 1) between the independent and correlated models reveals interesting differences between the two models across the four possible consideration segments (both considered, bus only considered, train only considered, and neither considered). The independent model suggests that the probability of consideration of both bus and train is less than the probability of consideration for bus alone, whereas the correlated model suggests no noticeable difference. The difference between the average probability of considering neither relative to the probability of considering train only is larger in the correlated model than the independent model. Thus, neglecting correlation in consideration of alternatives can lead to erroneous forecasts and misleading policy evaluations.

Key Factors Influencing Consideration of Bus and Train

The results presented in Table 3 (model M1b) provide important insights into the differences between the factors affecting consideration of bus and train for the work commute. Captivity status affects both bus and train consideration but in different ways. Captive workers (those without vehicle or driving knowledge) are more likely to consider bus than choice segment, but there is no discernible effect on train consideration. Workers belonging to the semi-captive segment, on the other hand, also have a greater tendency to consider both, but with a smaller preference than the captive segment.

Access and egress distances were found to affect only the consideration of train. The results show that workers whose residence (work) is within 1.2 km (1.0 km) from a railway station are more likely to consider train. Furthermore, the sensitivity of egress is 0.40 compared to 0.21 for access, suggesting that last mile connectivity has possibly a larger influence on consideration than the first-mile connectivity. On the other hand, the effect of access and egress is insignificant on consideration of bus. This may be due to the better accessibility and network coverage by bus than train. For instance, bus stops are usually more accessible than train stations in Chennai city (the average distance and coefficient of variation for bus stop distances were 0.53 km and 0.75, respectively for bus and 3.1 km and 1.2 for train, respectively).

Among the policy-sensitive factors, direct bus service increases the propensity to consider bus, indicating the burden of transfers on consideration and subsequently choice. Provision of information regarding bus timings and routes has a positive effect on the consideration of bus suggesting

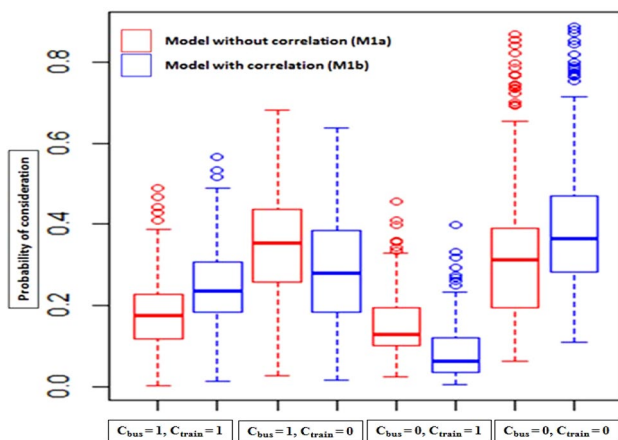


Fig. 1 Comparison of the probability of consideration of bus and train between model M1a and M1b

its role in attracting (unfamiliar) users to transit. Workers who perceived walking and crossing roads to the railway station as easy were more likely to consider train indicating that the influence of the quality of pedestrian infrastructure on considering train.

Amongst the contextual factors, workers who return home for lunch were less likely to consider bus or train for work commute. Such an activity may be possible for very short work distances where public transit is not attractive. For example, for work distances less than 2 km, consideration of both train and bus is lower. Furthermore, the time constraints of such a return home activity may also favor the use of personal modes. These findings show that contextual factors such as work distance and activity participation may play an important role on the consideration of public transit.

Role of Consideration of Public Transit on Primary Mode Choice for Work Trip

The results from models M2, M3, and M4 (presented in Table 4) are compared to understand the role of consideration of transit on the primary mode choice. Models M3 and M4 which include consideration effects clearly outperform Model M2 by 53 and 61 points, respectively, and the differences are significant at the usual 5% level.

The statistical inferences also are affected by neglecting the effect of consideration. For example, model M2 overestimates the magnitude of the alternate specific constants for bus and train compared to M3 and M4. Travel costs for auto and shared auto are insignificant in M2 (at 10% significance) but not in M3 and M4. In contrast, low income (less than Rs. 20,000) is significant in M2 but not in M4. On the other hand, a significant effect of high income (greater than Rs. 40,000) variable on bus consideration was found in model M4. Thus, neglecting consideration information could result in a reduction in the goodness-of-fit, possible bias, and misleading inferences.

The coefficients of the degree of consideration of bus and train in M3 and M4 were significant and positive indicating their positive role in mode choice. Also, the statistical test that the parameter for consideration probability on choice utilities is equal to 1 is rejected at the 5% significance level. If models M3 and M4 were constrained to have unit coefficients for log (consideration probabilities), the resulting goodness-of-fit is worsened by 31 and 3 points, respectively. Therefore, the Chi-square test also rejects the assumption that constrains the IAP coefficients to one.

The log-likelihood of model M4 is lower than M3 by 8 points. The Chi-squared value is 15.98 which is larger than the critical Chi-square value (with 5 degrees of freedom) of 11.07, and thus, M4 is significantly better than M3 at 5% significance level. Thus, the hypothesis of full mediation by consideration-related factors is to be rejected in favor

of partial mediation. However, the magnitudes of the log (consideration probability) coefficients were larger in M3 as compared to M4. This suggests possible overestimation in M3 due to the omission of some variables which not only affect choice through consideration indirectly but also directly.

Factors Influencing Primary Mode Choice for Work

The results from model M4 in Table 4 are discussed below. The factors influencing mode choice can be distinguished as fully mediated, partially mediated, or unmediated through consideration. If all variables that influence consideration indirectly affect choice only through consideration, then their effect is said to be ‘fully’ mediated. If in addition, a direct effect is observed for some variables after consideration effect is captured, these variables are only ‘partially’ mediated by consideration. Unmediated variables are those that do not affect consideration but directly influence choice utility. Differentiating these variables can aid in selecting appropriate policies to enhance consideration of transit and those that augment ridership.

The effect of captivity status and the contextual variable of returning home for lunch on bus choice were fully mediated by the consideration of bus. This implies that while captive users are more likely to consider bus, they were not significantly more likely to select bus once the consideration effect is captured. A similar interpretation also applies for workers who return home for lunch but in the negative direction. Those who return home for lunch are less likely to consider bus, but it has no additional influence on choice. In contrast, most variables (except semi-captive status) that influence train consideration were fully mediated. Access and egress distance to nearest railway stations had a negative influence on choice only through consideration probabilities.

The following variables showed the effect of partial mediation via consideration on the choice of bus. Transit service-related factors such as availability of direct bus to work and provision of accurate information regarding bus routes and timings indirectly increase the choice probability by increasing consideration, but also its likelihood of choice as the primary mode given consideration. Workers whose distance to work is less than 2 km have a lesser propensity to choose bus even when considered, and may represent cases where other non-motorized or IPT modes may be feasible or convenient. Workers belonging to the semi-captive segment were more likely to use both train and bus as the primary mode in addition to the positive influence on consideration noted in the section “[Key Factors Influencing consideration of Bus and Train](#)”. Workers with high income (> Rs. 40,000/month) are less likely to both consider and use bus as the primary mode.

Among the variables that affect the mode choice directly and not via consideration, travel time, and travel cost per kilometer negatively influence the choice of mode as expected. The sensitivity to travel time varies across modes and are highest for non-motorized modes (− 0.08) and auto-rickshaws (− 0.13). The sensitivity of time for personal vehicles is (− 0.04) which is larger than shared modes (bus − 0.02, train − 0.01, and share auto − 0.03). The lower sensitivity in shared modes may be due to the presence of considerable out-of-vehicle time components in travel time. In contrast, workers seem to be more sensitive to travel cost/km for car (− 0.93) which has a higher operational cost than other modes. The cost coefficient magnitude is larger and significant for transit and company bus (− 0.12) but insignificant for two-wheeler. This may explain why transit fare increases has led to lower transit ridership, whereas petrol price change has no effect on the decrease in two-wheeler shares. The insignificance of two-wheeler cost may suggest that its fuel cost is not perceived as much as an out-of-pocket cost unlike transit costs.

With regard to socio-demographic factors, income, number of two-wheelers, and gender influence the choice of mode. Workers with low income are less likely to choose car or auto as the primary mode. Female users are more likely to choose public transit modes than males, possibly because of the lack of driving knowledge among a greater fraction of female workers. Working women also have a greater preference for IPT modes (auto, shared auto, and company bus) than public transit, which may indicate the greater valuation of privacy and safety in these modes.

Implications Due to Policy and Change in Share of Consideration

The models estimated from this study can be used to predict the impact of some key policy-responsive and socio-demographic variables on both consideration and choice of public transit. Three different illustrative scenarios are analyzed in this context.

The first scenario examines the effect of the increase in vehicle ownership which can lead to a reduction in captive and semi-captive segments. A 5% decrease each in captive and semi-captive segments are considered. The reduction in captive segment is evenly distributed to the other two segments, whereas the workers who move from semi-captive group are assumed to move into the choice segment. The second scenario evaluates the role of improving access to railway stations. The proportion of users with access to railway station (within 1.2 km) is assumed to increase from 29% in the sample to 49%, perhaps as the growing network connectivity from metro services in the future. The third scenario investigates the role of more seamless transit operations. It is assumed that direct bus connectivity is enhanced from the current level of 60% (in the sample) to 80% in the future. The effect of these scenarios is modeled using the proposed model (M1b with M4) as well as other benchmark models discussed previously.

Table 5 presents the change in the share of the various choice sets with respect to the estimates from model M1b for each scenario. Overall, as expected, the model shows that a reduction in the share of captive and semi-captive users reduces the consideration of bus. However, there is an increase in the share of workers considering only train, indicating that an increase in the vehicle ownership might not necessarily decrease the consideration of all transit modes.

Table 5 Impact of scenarios on consideration of bus and train

Policy scenario	Model	$C_{bus}=1, C_{train}=1$ (%)	$C_{bus}=1, C_{train}=0$ (%)	$C_{bus}=0, C_{train}=1$ (%)	$C_{bus}=0, C_{train}=0$ (%)
Scenario 1: vehicle ownership increase	M1b	− 2.12	− 2.77	2.83	2.71
Scenario 2: increase in train access	M1b	3.68	− 3.16	6.34	− 1.35
Scenario 3: increase in direct bus services	M1b	3.02	5.75	− 8.78	− 4.15

Table 6 Impact of scenarios on choice of bus and train

Policy scenario	Bus			Train		
	Model M2 (%)	Model M3 (%)	Model M4 (%)	Model M2 (%)	Model M3 (%)	Model M4 (%)
Scenario 1: vehicle ownership increase	0.00	− 6.86	− 6.75	0.00	1.30	− 1.01
Scenario 2: increase in train access	0.00	− 0.73	− 0.75	0.00	7.16	6.57
Scenario 3: increase in direct bus services	0.00	8.19	8.25	0.00	− 3.42	− 3.66

An increase in the accessibility to railway station shows an increase in the share of workers considering train. This increase is mainly associated with the reduction in the share of workers considering only bus than those who consider neither of the transit modes. On the contrary, provision of direct bus services exhibits a decrease in the share of workers considering train as these workers now start considering bus more than train.

The impact of each scenario on the choice of bus and train as primary mode estimated from models M2, M3 and M4 by calculating the change in the probability of choosing bus and train compared to the base probabilities (shown in Table 6).

As Model M2 assumes that consideration has no effect on choice, there is no change in probabilities across all three scenarios. The model M3 overstates the change in bus shares for the first scenario relative to M4. Furthermore, model M4 predicts a decrease in train share with change in captivity, whereas M3 predicts an increase. This is because of the relative decrease in the share of semi-captive users (by 2.5%) compared to the original share.

Thus, the analysis of various scenarios demonstrates that the proposed model is able to differentiate the effect of alternative policies on consideration and choice.

Summary and Conclusions

In this study, joint discrete choice models have been developed for the consideration of public transit and primary mode choice for the work commute. The analysis is based on a dataset from a sample of workers in the Greater Chennai Metropolitan Area.

The following salient findings are observed in this study. Consideration probabilities of bus and train are positively correlated. Modeling consideration without accounting for the interrelationship reduces the goodness of fit and could possibly produce biased estimates. Some factors that influence choice of transit as the primary mode are partially mediated by consideration, whereas others are fully mediated by consideration. Neglecting this difference can lead to poor fit and erroneous policy evaluations. Captivity status influences consideration of bus differently from that of train. While captive and semi-captive users tend to consider bus, only semi-captive users were found to have a higher propensity to consider and choose train. Accessibility to train station near home and work had a contrasting effect on the consideration and choice of bus versus train. Although workers who have better access to train stations tend to consider train, they are also less likely to consider bus. A similar but smaller effect was also observed on the choice of bus and train. The effect of variables such as income, semi-captive segment, and service-related factors such as direct bus and information regarding bus timings on choice were partially

mediated by consideration of bus. Similarly, semi-captivity status partially mediates the relation between consideration and choice of train. Contextual factors and ease of last-mile connectivity were found to significantly affect mainly public transit consideration, indicating that consideration is influenced by overall activity and travel patterns rather than just work commute. These factors could provide a better understanding of consideration and choice of public transit, particularly in developing countries.

This study provides an understanding of some key factors influencing consideration and choice of public transit based on data from a sample of workers in Chennai city. The results from the study may be corroborated using data on work mode choice in other Indian cities. The key factors influencing consideration and choice of other segments such as non-workers and students remain to be investigated. The methodology developed in this study can be used in understanding potential segments where consideration probabilities can be increased in order to develop suitable policies for these segments. The primary scope of this study was on consideration of public transport, which can be expanded to account for the consideration of other modes in the future. Finally, models that allow for unobserved correlation across consideration and conditional choice can be investigated in future research.

Acknowledgements The authors acknowledge the opportunity provided by the 4th Conference of the Transportation Research Group of India (4th CTRG) held at IIT Bombay, Mumbai, India, in Dec. 2017. This work is supported through the Centre of Excellence in Urban Transport at IIT Madras, funded by the MoUD, GoI. Their support is gratefully acknowledged. The authors are also grateful to the anonymous reviewers for their constructive comments and suggestions on the original manuscript.

Funding This study was made possible partly due to funds made available through the research project sponsored by Ministry of Urban Development, Govt. of India.

References

1. Shocker AD, Ben-akiva M, Boccara B, Nedungadi P, Letters SM, Making CD, Behavior C, Ben-akiva M (1991) Consideration set influences on consumer decision-making and choice : issues, models, and suggestions. *Market Lett* 2(3):181–197
2. Nedungadi P (1990) Recall and consumer consideration sets: influencing choice without altering brand evaluations. *J Consum Res* 17(3):263–276. <https://doi.org/10.1086/208556>
3. Manski CF (1977) The structure of random utility models. *Theor Decis* 8(3):229–254. <https://doi.org/10.1007/BF00133443>
4. Cascetta E, Papola A (2001) Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. *Transp Res Part C Emerg Technol* 9(4):249–263. [https://doi.org/10.1016/S0968-090X\(00\)00036-X](https://doi.org/10.1016/S0968-090X(00)00036-X)
5. Gadepalli R, Tiwari G, Bolia N (2018) Role of user's socio-economic and travel characteristics in mode choice between city bus and informal transit services: lessons from household surveys in

- Visakhapatnam, India. *J Transp Geogr*. <https://doi.org/10.1016/j.jtrangeo.2018.08.017>
6. Bhat CR (1997) Work travel mode choice and number of non-work commute stops. *Transp Res Part B Methodol* 31(1):41–54. [https://doi.org/10.1016/S0191-2615\(96\)00016-1](https://doi.org/10.1016/S0191-2615(96)00016-1)
 7. Portoghese A, Spissu E, Bhat CR, Eluru N, Meloni I (2011) A copula-based joint model of commute mode choice and number of non-work stops during the commute. *Int J Transp Econ* 38(3):337–362. <https://doi.org/10.1007/s11116-007-9153-5>
 8. Paleti R (2015) Implicit choice set generation in discrete choice models: application to household auto ownership decisions. *Transp Res Part B Methodol* 80:132–149. <https://doi.org/10.1016/j.trb.2015.06.015>
 9. Swait J, Ben-Akiva ME (1987) Incorporating random constraints in discrete choice set generation. *Transp Res Part B* 21B(2):91–102
 10. Calastri C, Hess S, Choudhury C, Daly A, Gabrielli L (2019) Mode choice with latent availability and consideration: theory and a case study. *Transp Res Part B Methodol* 123:374–385. <https://doi.org/10.1016/j.trb.2017.06.016>
 11. Enam A (2010) Developing a comprehensive mode choice model to capture the preferences for mass rapid transit in Dhaka. Bangladesh University of Engineering and Technology, Bangladesh
 12. Mark JIG, Marcel GD (1979) The dogit model. *Transp Res Part B Methodol* 13B(2):105–111
 13. da Sanches SP, de Arruda FS (2002) Incorporating nonmotorized modes in a mode choice model. *Transp Res Record J Transp Res Board* 1818(1):89–93. <https://doi.org/10.3141/1818-14>
 14. Srinivasan KK, Pradhan GN, Naidu GM (2007) Commute mode choice in a developing country. *Transp Res Record J Transp Res Board* 2038(1):53–61. <https://doi.org/10.3141/2038-07>
 15. Rashedi Z, Habib KN (2019) A semi-compensatory choice model with probabilistic choice set: combining implicit choice set within probabilistic choice set formation. In: 98th Annual Meeting of TRB. Washington DC, p 25
 16. Kunhikrishnan P, Srinivasan KK (2017) Choice set variability and contextual heterogeneity in work trip mode choice in Chennai City. *Transp Lett* 11(4):174–189. <https://doi.org/10.1080/19427867.2017.1299396>
 17. Swait J (2001) A non-compensatory choice model incorporating attribute cutoffs. *Transp Res Part B Methodol* 35(10):903–928. [https://doi.org/10.1016/S0191-2615\(00\)00030-8](https://doi.org/10.1016/S0191-2615(00)00030-8)
 18. Castro M, Martínez F, Munizaga MA (2013) Estimation of a constrained multinomial logit model. *Transportation* 40(3):563–581. <https://doi.org/10.1007/s11116-012-9435-4>
 19. Martínez F, Aguila F, Hurtubia R (2009) The constrained multinomial logit: a semi-compensatory choice model. *Transp Res Part B Methodol* 43(3):365–377. <https://doi.org/10.1016/j.trb.2008.06.006>
 20. Sana B, Outwater M, Ferdous N, Woodford B, Lobb J (2011) Characteristics of premium transit services that affect choice of mode. The National Academies Press, Washington
 21. Tirachini A, Hensher DA, Rose JM (2013) Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. *Transp Res Part A Policy Pract* 53:36–52. <https://doi.org/10.1016/j.tra.2013.06.005>
 22. Manoj K, Vikas A (2016) Public Transport service quality and passenger satisfaction: a case of UPSRTC, Agra, India. *Pacif Bus Rev Int* 8(11):82–92
 23. Srinivasan S, Rogers P (2005) Travel behavior of low-income residents: studying two contrasting locations in the city of Chennai, India. *J Transp Geogr* 13(3):265–274. <https://doi.org/10.1016/j.jtrangeo.2004.07.008>
 24. Srinivasan KK, Bhargavi P (2007) Longer-term changes in mode choice decisions in Chennai: a comparison between cross-sectional and dynamic models. *Transportation* 34(3):355–374. <https://doi.org/10.1007/s11116-007-9116-x>
 25. Suman HK, Bolia NB, Tiwari G (2017) Comparing public bus transport service attributes in Delhi and Mumbai: policy implications for improving bus services in Delhi. *Transp Policy* 56(March):63–74. <https://doi.org/10.1016/j.tranpol.2017.03.002>
 26. Advani M, Tiwari G (2005) Evaluation of public transport systems: case study of Delhi Metro. In: START-2005 Conference held at IIT Kharagpur, India, No. 2003, pp 1–8
 27. Maitra B, Dandapat S, Chintakayala P (2015) Differences between the perceptions of captive and choice riders toward bus service attributes and the need for segmentation of bus services in Urban India. *J Urban Plan Dev* 141(2):04014018. [https://doi.org/10.1061/\(ASCE\)UP.1943-5444.0000205](https://doi.org/10.1061/(ASCE)UP.1943-5444.0000205)
 28. Das S, Pandit D (2015) Determination of level-of-service scale values for quantitative bus transit service attributes based on user perception. *Transportmetr A Transp Sci* 11(1):1–21. <https://doi.org/10.1080/23249935.2014.910563>
 29. Tiwari G, Jain D, Rao KR (2016) Impact of public transport and non-motorized transport infrastructure on travel mode shares, energy, emissions and safety: case of Indian cities. *Transp Res Part D Transp Environ* 44:277–291. <https://doi.org/10.1016/j.trd.2015.11.004>
 30. Ashalatha R, Manju VS, Zacharia AB (2012) mode choice behavior of commuters in Thiruvananthapuram City. *J Transp Eng* 139(5):494–502. [https://doi.org/10.1061/\(asce\)te.1943-5436.0000533](https://doi.org/10.1061/(asce)te.1943-5436.0000533)
 31. Das S, Pandit D (2016) Methodology to determine service delivery levels for public transportation. *Transp Plan Technol* 39(2):195–217. <https://doi.org/10.1080/03081060.2015.1127541>
 32. Schwanen T, Mokhtarian PL (2005) What affects commute mode choice: neighborhood physical structure or preferences toward neighborhoods? *J Transp Geogr* 13(1):83–99. <https://doi.org/10.1016/j.jtrangeo.2004.11.001>
 33. Office of the Registrar General & Census Commissioner, India. 1–2. <https://censusindia.gov.in/2011census/dchb/Tamilnadu.html>. Accessed 23 May 2019
 34. Associates W (2010) Chennai Comprehensive Transportation Study. Chennai, India

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.