Heterogeneous Valuation of Quality Dimensions of Railway Freight Service by Chinese Shippers

Choice-Based Conjoint Analysis

Duan, L.; Rezaei, Jafar; Tavasszy, Lóri; Chorus, Caspar

DOI
10.3141/2546-02

Publication date
2016

Document Version
Final published version

Published in
Transportation Research Record

Heterogeneous Valuation of Quality Dimensions of Railway Freight Service by Chinese Shippers

Choice-Based Conjoint Analysis

Liwei Duan, Jafar Rezaei, Lóránt Tavasszy, and Caspar Chorus

This paper operationalizes and tests approaches to identify market segments for rail freight services and measures the importance that customers attach to rail service attributes (i.e., transport cost, time, frequency, reliability, and safety). The approach is based on choice-based conjoint analysis in which heterogeneity is captured by means of latent-class analysis. The research is novel in several respects. First, it addresses the diverse valuation of service preferences by shippers who use rail transport. Second, besides estimates for rail users who contract for transport services, the analysis also arrives at new estimates for forwarders as immediate clients for rail services. Third, in addition to the conventional random utility maximization (RUM) model, the paper discusses trials with a random regret minimization (RRM) model and a hybrid RUM-RRM model. Finally, the research produces unique values for China, one of the largest rail transport markets in the world.

The freight transportation industry is the backbone for domestic and international trades, with improvements in the freight transportation system invariably finding their way into increased trade and economic growth. Identification of potential improvements relies on a thorough knowledge of how users perceive the quality of transport services and what they think are important attributes. Usually, these perceptions and preferences will vary among types of users (e.g., shippers, forwarders) and types of goods (e.g., containerized, bulk).

Rail transport has great advantages for long-distance and heavy shipments, which are salient features for large regions such as the United States, Europe, and China. Its disadvantages are also obvious: railways are less flexible and diversified when compared with road and air freight transport. These disadvantages become problematic when shippers are willing to pay more for higher-quality freight transport. If rail transport cannot keep up with changes in demand, shippers might switch to other modes. Thus, shippers' preferences should be studied, not only about their mode preferences but about their preferences within modes.

L. Duan, School of Transportation and Logistics, Southwest Jiaotong University, No. 111, Beiyiduan, 2nd Ring Road, Chengdu 610031, Sichuan, China. J. Rezaei, L. Tavasszy, and C. Chorus, Faculty of Technology, Policy, and Management, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, Netherlands. Corresponding author: L. Tavasszy, l.a.tavasszy@tudelft.nl.

Transportation Research Record: Journal of the Transportation Research Board, No. 2546, Transportation Research Board, Washington, D.C., 2016, pp. 9–16.
DOI: 10.3141/2546-02

This research need becomes further emphasized in light of the observation that, in the past few decades, the competitiveness of rail freight services has been declining. And railways have been struggling to maintain their market share. The market share of China's railway freight transport has decreased to 8.7% (in tons) and 14.9% (in ton-kilometers) in 2014 compared with 14.5% (in tons) and 25.9% (in ton-kilometers) in 2005 (*1*). Therefore, understanding the requirements of shippers toward rail services is of vital importance to rail carriers. As preferences in the market are likely to be quite diverse, the study of variations in preferences among shippers becomes more important. The objective of this paper is to improve understanding about this heterogeneity in service valuation for the market of rail transport.

Ample literature discusses both service preferences and service design, which are rooted in various branches of academic literature, including operations, behavioral economics, and marketing research. Recent work on service design that uses knowledge of the differentiated characteristics of shippers includes several sources (*2–6*). The focus in this paper is on the demand (as opposed to the design) side of the market, however. Freight service preferences between service modes have been studied since the 1970s (*7–11*). To the authors' knowledge, however, no literature has explored the demand preferences of shippers who use railways services, particularly that with a focus on the heterogeneity of preferences. Even though within-mode preference choice was included sporadically in papers, those studies did not present estimates specifically focused on railway shippers (*12–17*). Halse and Killi investigated customers' valuation of time and reliability in rail freight of Norway but mainly forwarders (*18*). To bridge this gap in the literature of demand and preference studies for railway freight services, the current authors developed and applied a survey-based approach for the railway services market in China.

In sum, the main contribution of this study is as a within-mode study that addresses heterogeneity in preferences from the side of current rail service customers, in contrast with existing studies that estimate either average preferences for rail carriers or preferences across all modes of transport. The paper presents a case study for the Chinese rail freight market.

The outline of this paper is as follows. A brief review of the literature on relevant performance attributes, as seen by shippers, and on appropriate measurement methods comes next. That review is followed by a description of the design of the choice experiment as well as of the choice models estimated on the data. Then the results are reported; they are followed by, first, a market

segmentation analysis that uses a latent-class (LC) model and, finally, conclusions.

LITERATURE REVIEW

A shipper's demand for a mode of freight service depends on the levels of characteristics built into the service and the relative contributions of these characteristics to the shipper's production and distribution activities (*19*). Research on which service attributes are considered relevant by shippers had already been identified in the 1970s. McGinnis found that speed and reliability, rates, and loss and damage were the top three important groups of variables affecting shippers' choice (*7*). Lu listed time, frequency, delivery reliability, and the like as service attributes and implied that the timing-related services factor was the most important criterion for shippers (*20*). Danielis et al. selected cost, time, reliability, and damage as important attributes and concluded that shippers hold a strong preference for quality attributes over cost (*10*). A review by Cullinane and Toy showed that cost, speed, and reliability in transit time were the most preferred attributes by researchers (*9*). Feo-Valero et al. asserted that transport cost, time, reliability in delivery time, frequency, and reliability in delivery conditions were the attributes used most frequently (*21*).

Two possible ways to measure shippers' preferences on service attributes are revealed preference and stated preference (SP) studies (*22*). In the current case, SP was more appropriate than revealed preference when new service alternatives were introduced and no historical data could be obtained, as SP can elicit shippers' preferences by simulating shippers' choice behavior about hypothetical alternatives. In addition, because of SP's advantage in data acquisition, SP experiments are applied widely (*23, 24*) by freight transport researchers to analyze issues related to value of time (VOT) and value of service (VOS) (*14, 17, 25–29*).

VOT and VOS issues are important to the understanding of stakeholders' tradeoffs and utilities toward freight transport cost and time as well as to all essential service attributes. Vieira studied the influence of shippers' attitudes to service attributes and addressed the use of SP for market segmentation, albeit at the level of the entire freight transport market across all modes of transport (*25*). Other uses of VOT include infrastructure improvement, environment impact analysis, and revenue management (*20*). For these purposes, several studies used SP models for estimating the value of travel time savings (*14, 17, 26–28*). Danielis et al. (*10*) and Danielis and Marcucci (*30*) estimated logistics managers' preferences for road-only and non-road-only transport services. Bergantino and Bolis derived forwarders' values of time savings, frequency, and reliability for maritime services and road services through an adaptive SP experiment (*31*). Zamparini et al. determined the relative importance and monetary values attached to quality attributes of freight transport with an SP study (*32*). Arunotayanun and Polak estimated a model with heterogeneity in preferences by looking at the choices for road and rail services for specific commodity groups (*13*).

Despite the wide application of SP in freight transport, no specific study has appeared about shippers' demand preferences for railway-only freight services, especially for China, the biggest developing country in the world. In addition, heterogeneity has hardly been addressed empirically. This dearth of research was confirmed by two systematic overviews by Zamparini and Reggiani (*33*) and Feo-Valero et al. (*21*) and by the current authors' additional search for later work.

CHOICE EXPERIMENT, DATA COLLECTION, AND CHOICE MODELS

Design of Choice Experiment

Choice-based conjoint (CBC) analysis, as one form of SP methodologies, was chosen to design the choice experiment in this study because of its advantages in dealing with the complexity arising with the growth of the number of attributes and levels (*34*). The CBC questionnaire consists of several choice tasks, and each one contains a set of multiple alternatives (profiles) that forms a combination of attribute levels (*35*).

On the basis of the existing studies reviewed earlier, the authors selected the five attributes of cost, time, frequency (per week), reliability (i.e., punctuality), and safety (no loss or damage), each with four levels, to specify the hypothetical choice alternatives (freight services). Table 1 reports the four levels of service for each attribute. For service reliability and safety, the numbers are the ratios of service not being delayed and cargo not being lost or damaged. For cost, time, and frequency, the numbers are percentage changes to absolute values. The reason for this measurement is that absolute values for these attributes vary a lot because of the wide ranges of transport distances (from hundreds to thousands of kilometers).

Profiles use orthogonal fractional factorial design to reduce respondent burden. Orthogonality ensures a systematic variation in attribute levels to arrive at a relatively efficient (i.e., with a limited number of respondents) data collection effort. This experiment was not constructed by using recent advances in efficient experimental design; in hindsight, especially in light of the fact that the sample was relatively small, application of such efficient design routines would have been useful, as it might have led to smaller standard errors.

In this case (with five attributes with four levels each), the orthogonal fractional factorial design resulted in 16 profiles. With the shifting method, these profiles were subsequently arranged in 16 choice sets containing four hypothetical choice alternatives each (*36*). To maximize the realism of the choice task, respondents were given the opportunity (at each choice task) to opt out and to choose the "[n]one of these" option. And that option was modeled through a separate utility function without taking attribute level changes into account. In addition, four general questions were asked to identify respondent characteristics. An example of the CBC questionnaire is shown in Table 1. The actual questionnaire is presented in Chinese, with the same choice situations presented in Table 1.

Data Collection

The target population for this survey consisted of shippers experienced in railway freight services in China. To allow segmentation, shippers were classified into three categories: individuals, companies, and freight forwarders.

The CBC study was conducted on one railway company in China, for which the authors selected and distributed questionnaires to 543 shippers randomly from that company's clients, of which 114 completed questionnaires were returned; thirty-one of those were excluded because of missing data. The 83 remaining responses resulted in 1,660 choice observations upon which the choice model was estimated; of these observations, the number of opt-out choices (Option 5 in Table 1) was 157. Of the 83 respondents, 45 nonforwarders (13 individual shippers and 32 company shippers) and 38 freight forwarder shippers were counted. By considering

TABLE 1 Example of CBC Choice Task

	If These Were Your Freight Service Options by Railway, Which Would You Choose?				
Attribute	Option 1 (%)	Option 2 (%)	Option 3 (%)	Option 4 (%)	Option 5
Cost	−5	+5	+10	−10	None of these
Time	−20	−10	+10	+20	
Frequency	+20	−20	−10	+10	
Reliability	95	100	70	85	
Safety	90	95	100	80	
Choice:					

containerization as another criterion, the respondents were categorized as 17 container shippers and 26 noncontainer shippers. The remaining 40 were both container and noncontainer shippers.

The authors do not claim that the sample is representative of the population of Chinese rail freight shippers and forwarders in relation to cargo type and other potentially relevant characteristics. In the absence of reliable data at the population level, and given that the study is aimed to be a first exploration of (heterogeneity in) freight transport preferences, the authors did not explore the representativeness issue here. However, many previous studies (see references at the beginning of the paper) have highlighted that the collection of a large and representative sample of freight shippers and transporters is very labor- and time-intensive—and often impossible, given conventional time and financial budgets available for this type of research. This complexity is widely considered to be an important reason why freight choice modeling efforts have been less widespread than their counterparts in a passenger mobility context.

Choice Models: Random Utility, Random Regret, and Hybrid Forms

The data collected by CBC experiments consists of a series of discrete choices between several hypothetical alternatives. From these choices, underlying preferences and tradeoffs can be inferred by estimating discrete choice models on the data. In the field of choice modeling, linear-in-parameter random utility maximization (RUM) models are by far the most widely adopted approach (*37*). These models assume that, when choosing between different options, decision makers attach a utility to each option and choose the one with the highest utility. This utility consists of the sum of an error term that represents incomplete knowledge on the part of the analyst and a linear-additive function of characteristics of the choice option and associated parameters (the latter representing tastes or decision weights associated with these characteristics). Different choice probability formulations can be obtained on the basis of the assumed distribution of the error term. The authors assumed that the errors were independent and identically distributed extreme-value Type I distributions across alternatives and choice tasks, with a variance that normalized to $\pi^2/6$; this calculation resulted in the well-known multinomial logit model (MNL) model, which is the workhorse model for discrete choice analysis.

More recently, several papers offered the random regret minimization (RRM) model as an alternative way of modeling choice behavior; examples include Chorus (*38*) and Chorus et al. (*39*). The RRM approach is based on the notion that, when choosing, people anticipate and aim to minimize regret rather than maximize utility. Regret arises when one or more nonchosen alternatives perform better than the chosen one in relation to one or more attributes. The RRM model postulates that a (further) deterioration of an attribute for which an alternative already has poor performance (relative to its competition) cannot be compensated for by an equally large improvement of another—equally important—attribute on which the alternative already has a strong performance (relative to its competition). A recent overview study showed that which of two models performs better in model fit and out-of-sample predictive ability is highly data set–specific. Equations 1 and 2 give the mathematical formulations of the RUM and RRM models, respectively.

$$U_i = V_i + \varepsilon_i = \sum_m \beta_m x_{im} + \varepsilon_i \tag{1}$$

$$RR_i = R_i + \upsilon_i = \sum_{j \neq i} \sum_m \ln\left(1 + \exp\left[\beta_m \left(x_{jm} - x_{jm}\right)\right]\right) + \upsilon_i \tag{2}$$

where

U_i, RR_i = random utility and regret, respectively, associated with considered alternative i;
V_i, R_i = observed or systematic utility and regret, respectively, associated with i;
ε_i, υ_i = for RUM error terms (and for RRM, their negatives) associated with i of independent and identically distributed extreme-value Type I;
β_m = estimable taste parameter associated with attribute x_m that differs between the RUM and RRM models in magnitude and behavioral interpretation; and
x_{im}, x_{jm} = values associated with attribute x_m for, respectively, the considered alternative i and another alternative j.

This paper explored both the RUM and RRM models and hybrid variants, not because the authors believe that one of the two models is potentially superior to the other from a theoretical or behavioral viewpoint, but simply with a view to being able to compare the two approaches empirically and to choose the specification that best fits the data.

Choice probabilities for the RUM and the RRM models, respectively, are as follows (both in MNL form):

$$P(i) = \frac{\exp(V_i)}{\sum_{j=1\ldots J} \exp(V_j)} \tag{3}$$

$$P(i) = \frac{\exp(-R_i)}{\sum_{j=1\ldots J} \exp(-R_j)} \tag{4}$$

Furthermore, on many occasions, a hybrid model (in which some attributes are processed by utility maximization and others by regret minimization) fits the choice data best (*39*). Behaviorally, such a hybrid model implies that some attributes are assumed by the researcher to be processed by the decision maker in a linear-additive RUM way while others are assumed to be processed in an RRM fashion. If Q attributes are being implemented outside the regret function and $M - Q$ inside the regret function [where $Q > 0$, $(M - Q) > 0$, and M = number of attributes], the systematic part of such a hybrid utility–regret function would look like this:

$$V_i = \sum_{m=1...Q} \beta_m x_{im} - \sum_{j \neq i} \sum_{m=Q+1...M} \ln\left(1 + \exp\left[\beta_m\left(x_{jm} - x_{jm}\right)\right]\right)$$

Choice probabilities are given by the RUM-based MNL formulation shown earlier.

In addition, to allow for distinction in attribute importance between different segments with observed heterogeneity, a base beta plus an additional segment-specific beta are used. Combined, these assumptions reduce to the following utility function (for succinctness, its regret-based and hybrid counterparts are not presented):

$$U_i = (\beta_C + \beta_{CM}) \times \text{cost}_i + (\beta_T + \beta_{TM}) \times \text{time}_i + (\beta_F + \beta_{FM}) \times \text{freq}_i + (\beta_R + \beta_{RM}) \times \text{reli}_i + (\beta_S + \beta_{SM}) \times \text{safe}_i + \varepsilon_i \quad (5)$$

where

cost_i, time_i, freq_i, reli_i, and safe_i = transport cost, time, frequency, reliability, and safety of service i, respectively;

β_C, β_T, β_F, β_R, and β_S = base coefficients to be estimated, with respect to the five service attributes mentioned above; and

β_{CM}, β_{TM}, β_{FM}, β_{RM}, and β_{SM} = segment-specific coefficients added to base coefficients.

In this case, reliability is defined as punctuality and modeled as the percentage of on-time deliveries, a common way to measure reliability in logistics. Meanwhile, the authors' approach to incorporating reliability into SP experiments and to capturing value of reliability (VOR) in choice models is only one of several possible approaches. Excellent, comprehensive reviews appear elsewhere (*40–43*). Exploration of different approaches to operationalize reliability in SP and to model VOR in choice models is left for future research because of space limitations.

MNL AND SEGMENTATION ANALYSIS

The first part of the empirical analyses here uses the linear-in-parameters RUM model (MNL form). NLOGIT software (*44*) and Latent GOLD software (*45*) were used for model estimation. Here, shippers handling both container and noncontainer cargoes are joined with noncontainer shippers so as to compare estimation results between different allocations. MNL models were estimated for the full sample and for each segment. Results appear in Table 2.

Clearly, for all shippers, every attribute coefficient has expected signs and is highly significant, both indications that shippers prefer lower cost, shorter time, and better service frequency, reliability, and safety.

For forwarders and nonforwarders, almost all coefficients are significant and have expected signs. Forwarders show less attention to cost, time, and reliability, but their valuation of safety is slightly higher than that of nonforwarders. (The segment-specific parameter for frequency is not significant). The results illustrate that forwarders would trade higher cost or time to more safety service.

The container and noncontainer segments show more strongly differentiated preferences about the choice of freight services. For the container segment, without consideration of cost and frequency because of their insignificance, safety is less valued than for the noncontainer segment; the container segment is more likely to choose freight services with shorter time and better punctuality.

In addition, because the parameters in this case are expressed as percentages of the absolute values of attributes, the VOT and VOR of all shippers could be roughly calculated for VOT by multiplying the ratio of the estimated time coefficient (0.0207) and, for VOR, by multiplying the reliability coefficient (0.0478), of the estimated

TABLE 2 Estimation Results with MNL Model

	Value by Shipper Type									
	All Shippers		Nonforwarder Shippers		Forwarder Shippers[a]		Noncontainer Shippers		Container Shippers	
Parameter	Coeff. (β_m)[a]	SE	Coeff. (β_{bm})	SE	Coeff. (β_{am})	SE	Coeff. (β_{bm})	SE	Coeff. (β_{am})	SE
Beta_cost	–0.0487***	0.0044	–0.0636***	0.0062	0.0309***	0.0089	–0.0520***	0.0050	0.0138	0.0110
Beta_time	–0.0207***	0.0022	–0.0248***	0.0030	0.0084*	0.0044	–0.0187***	0.0025	–0.0104*	0.0056
Beta_freq	0.0136***	0.0021	0.0116***	0.0029	0.0043	0.0043	0.0122***	0.0029	0.0083	0.0054
Beta_reli	0.0478***	0.0035	0.0563***	0.0045	–0.0164***	0.0056	0.0423***	0.0037	0.0316***	0.0073
Beta_safe	0.0994***	0.0055	0.0929***	0.0061	0.0157***	0.0056	0.1075***	0.0059	–0.0338***	0.0073
Beta_Opt_Out	13.0925***	0.6539	13.2487***	0.6601	na	na	13.298***	0.6632	na	na
Null LL	–2,137.3335		–2,137.3335				–2,137.3335			
Final LL	–1,680.3272		–1,667.4152				–1,664.7492			
Rho-squared	.2138		.2199				.2211			

NOTE: Coeff. = coefficient; SE = standard error; na = not applicable; LL = log likelihood.
[a]As indicated in the main text, for forwarder shippers, the column coeff. (β_{am}) lists the additional coefficients β_{am} that should be added to the base coefficient coeff. (β_{bm}) when respondents are classified into the forwarder category. The same applies to container shippers.
*** = significant at 1% level; ** = significant at 5% level; * = significant at 10% level.

cost coefficient (0.0487) by the transport cost per hour for train services, as discussed in De Jong et al. (*15*). And the transport cost per hour for railway companies in China could be derived by multiplying the average tariff of both container and noncontainer cargoes [0.1044 Chinese yuan (CNY; 1 CNY = US$0.15 in January 2016) per ton-kilogram (*46*)] by the average travel speed, 34 km/h (*47*). Thus, VOT and VOR for all shippers are 1.51/ton-h and CNY3.48/ton-h, respectively. And the VOTs in this case are smaller than those in De Jong et al. (*14*) but within the range of VOTs from international literature reviewed by Feo-Valero et al. (*21*).

LC ANALYSIS

To study whether (and if so, in what ways) additional unobserved heterogeneity in shippers' preferences is not taken into account by segmentation that is based on observable characteristics, a series of so-called LC models was estimated. These models assume the existence of LCs (as opposed to observable segments) of respondents, each with its own set of preferences. Each individual is assigned to a different class, up to a probability. In other words, preferences (are assumed to) vary across but not within classes. As argued in, for example, Greene and Henshers, the LC approach offers a flexible way to capture preference heterogeneity, which is becoming increasingly popular both within the field of transportation and in adjacent fields (*47*). Arunotayanun and Polak estimated an LC model in a freight transport context and concluded that it provided a better explanation of heterogeneity than segmentation into commodity groups (*13*). They did not explore further classification possibilities to explain heterogeneity. With the data set used here, the authors were able to compare the results of classifying users by type of shipper and type of load with those of an LC model.

This study specified the number of classes (ranging from one to seven) and assessed which number gives the best model fit (after correcting for losses in degrees of freedom). Furthermore, the LC approach was combined with RUM, RRM, and hybrid RUM-RRM models, in the sense that a RUM, RRM, or hybrid choice model was assumed for all classes and model fit differences were compared across decision rules. Decision rule refers to the heuristic embedded in the choice model; in this case, the heuristic can be RUM, RRM, or hybrid RUM-RRM (see earlier equations). The authors' models assume heterogeneity in preferences across classes but do not allow for potential heterogeneity in decision rules across classes. After preliminary analysis in which all possible combinations of decision rules across attributes were explored, the best fitting (in relation to final log likelihood) hybrid model was found to treat reliability and safety (and the opt-out constant) as utility maximization attributes and cost, time, and frequency as regret minimization attributes. On the basis of the results obtained in preliminary analyses and to keep model tractability and interpretability at reasonable levels, the authors chose not to allow decision rules to differ across classes. For examples of papers in contexts other than freight transport in which allowing for decision rule heterogeneity led to substantial gains in model fit, see Boeri et al. (*48*), Hess and Chorus (*49*), and Hess et al. (*50*).

Individuals were assigned to LCs on the basis of a so-called membership function. This function may use a variety of covariates to predict to which class an individual's choices belong. This membership function features error terms and accounts for limited information on the part of the analyst; such error terms imply that class membership is predicted to a probability (in a way that is very similar to how choice predicts the probability that a certain alternative is chosen). This study tried several specifications for membership functions, by basing them on whether the respondent was a forwarder and on whether the respondent carried containers. Estimation results suggested that these factors played no role in determining class membership probabilities; this absence of significant covariates in the membership function implies that, apparently, the heterogeneity in preferences that comes to the fore in the current LC analyses is of a different type than that of the segment-related heterogeneity discussed earlier.

That LC estimation algorithms may get stuck in local (as opposed to global) maxima, especially when the number of classes grows, is well known. This weakness is one reason that the current study was limited to a maximum of seven classes. The fact that the improvement in model fit (corrected for losses in degrees of freedom) declines quickly with the number of classes beyond seven is another reason. To test for the potential presence of local maxima, the authors reestimated each model run by using five starting values that were generated by the software. When different estimation results were obtained for one or more of these five runs, the best-fitting model was reported and an underline added in Table 3 to signal that this result may have been a local maximum. The Bayesian information criterion (BIC) was used to assess model fit; this criterion penalizes losses in degrees of freedom and as such is a more conservative measure of model fit than the final log likelihood (and more conservative in this regard than the often-used Akaike information criterion); a lower BIC signals a better model fit.

Model fit results for each number of classes and for each decision rule (RUM versus RRM versus hybrid) are presented in Table 3. A few results catch the eye: first, capturing preference heterogeneity is beneficial in the sense that large improvements in model fit are obtained with a shift from a single-class to a multiple-class model. The largest gains are achieved with a shift from one to two classes, with the size of improvements generally decreasing as more LCs are added. Second, while the relative superiority of RUM and RRM depends on the number of classes, the hybrid model performs better than RUM and RRM for nearly every number of classes (except for four classes). In sum, considerable heterogeneity is found in preferences (across individuals) as well as in decision rules (across attributes).

Taking into account that model fit improvements become much less pronounced when the number of classes exceeds three, and for

TABLE 3 BIC Values of RUM, RRM, and Hybrid RUM-RRM Models

	Model Fit Result by Decision Rule		
Number of Classes (parameters)[a]	RUM	RRM	Hybrid RUM-RRM
1 (6)	3,387	3,392	3,384
2 (13)	3,055	3,075	3,047
3 (20)	2,910	2,925	2,900
4 (27)	2,837	2,912	2,839
5 (34)	2,847	2,862	2,814
6 (41)	2,838	2,828	2,818
7 (48)	2,839	2,824	2,792

[a]Six parameters per class; one membership constant per class, except for the one-class model.

TABLE 4 Coefficient Estimates for Service Attributes: Three-Class Model

Class	Parameters	Cost	Time	Frequency	Reliability	Safety	Size (%)
1	Coeff.	0.0006	–0.0008	–0.0000	–0.0147**	0.0242**	18.43
	SE	0.0039	0.0017	0.0019	0.0063	0.0099	
	p-value	.88	.66	.99	.020	.015	
2	Coeff.	–0.0392***	–0.0031	0.0041**	0.0706***	0.2672***	41.18
	SE	0.0049	0.0019	0.0017	0.0071	0.0264	
	p-value	2.1 E–15	.11	.017	1.9 E–23	4.4 E–24	
3	Coeff.	–0.0343***	–0.0245***	0.0166***	0.1196***	0.0949***	40.38
	SE	0.0038	0.0031	0.0021	0.0132	0.0125	
	p-value	1.6 E–19	3.8 E–15	4.2 E–15	1.2 E–19	3.1 E–14	

*** = significant at 1% level; ** = significant at 5% level; * = significant at 10% level.

reasons of interpretability, the authors decided to focus on the three-class model (hybrid RUM-RRM) for the remainder of the paper (Table 4).

Most parameters are significant at reasonable levels. In Table 4, almost all attributes for all classes have expected signs, except for the smallest (Class 1), in which the parameter for reliability is negative, which indicates that shippers in Class 1 are reliability averse. This counterintuitive sign for reliability appears in every model (i.e., RUM, RRM, hybrid; from one to seven classes) and always in the smallest class. No obvious reason explains this result. Different valuations in parameters of the five attributes can be readily observed across the three classes. Shippers in Class 1 are less sensitive to all attributes except service safety. Class 2 consists of shippers who are highly sensitive to service safety, reliability, and cost, while shippers in Class 3 show stronger preferences toward reliability than safety and cost. Service time and frequency seem less important for shippers in all classes.

One additional comment on methodology is required before the conclusions: this study made no attempt at validating the model (except for exploring face validity by inspecting parameter signs). An often-used validation approach for the type of models applied here is to use part of the data (e.g., 2/3) for estimation and the remaining part for out-of-sample validation. For this validation sample, the estimated model is applied to compute choice probabilities for chosen alternatives so as subsequently to obtain fit measures such as rho-squared or hit rate (percentage correctly predicted). Such validation is generally considered a useful test against overfitting on the estimation data. The reason that this study did not embark on such validation exercises is that the sample was already rather small (containing 83 valid respondents) compared with conventional sample sizes in SP experiments in passenger transport. Particularly because the final LC models consumed a considerable number of parameters, the authors found that partitioning the data into even smaller estimation and validation samples seemed to be unwise; doing so would have left too few degrees of freedom for reliable model estimation.

CONCLUSIONS

This paper elaborated shippers' preferences on railway freight services in China, and those preferences could be further compared with those for other modes in other countries or regions in relation to the importance of attributes and other primary features. The survey proved valuable for analyzing shippers' preferences, including heterogeneity issues.

Through the CBC study, the authors found that shippers make trade-offs when facing a series of freight services, especially when the services are highlighted by a bundle of attributes such as cost, time, frequency, reliability, and safety. Therefore, marketing on one attribute alone will be unfavorable for railway companies. In the case in this paper, besides paying attention to costs, shippers also consider transport time, reliability, and safety issues, findings that align with the consensus in the broader literature on freight mode choice that quality attributes play an important role. The ongoing development of the service industry has spawned customers' derived need for service quality, especially for those with high-value cargoes. This finding indicates that pricing may not be the best way to make railway freight services more attractive. In contrast, multidimensional marketing that focuses on both service quality and price may allow railways to gain more popularity among shippers. For the purpose of building robust forecasts, future research could address the addition of revealed preference data and variations in the choice of service attributes.

This paper examined heterogeneous preferences and confirmed them with the RUM, RRM, and hybrid RUM-RRM models. Generally, according to the BIC values in this case, neither RUM nor RRM is the clear winner; however, the hybrid RUM-RRM model shows better permanence in BIC than the RUM and RRM models for all samples (expect for one class for which the hybrid model performed equally well as the best RUM model). This result should provide a good start for introducing the hybrid RUM-RRM model into choice behavior studies in freight services. Because shippers' choice behavior is complicated and affected by multiple variables, consideration of the rules of both utility maximization and regret minimization can improve estimation accuracy and significance. As a referee of this paper noted, the promising performance of the hybrid RUM-RRM model compared with a presumably more rational RUM model may come as a surprise because decision makers in a competitive business environment are generally considered likely to exhibit rational behavior. Nonetheless, a recent study of the behavior of Swiss logistics managers also observed a strong performance of the RRM model compared with the RUM model (*51*). Exploration of the sources for this seemingly irrational behavior among highly trained professionals in a highly competitive industry is another interesting avenue for further research.

Further empirical analysis with segmentation of Chinese railway shippers by hybrid RUM-RRM model revealed that heterogeneous valuation of railway freight service quality dimension exists among shippers. As shippers continue to increase emphasis on service experiences and preferences, service providers should adopt as a new

dominant marketing strategy the offering of differentiated freight services to those with heterogeneous preferences. Price-sensitive shippers could receive offers of services at lower rates and an acceptable service level. And quality-sensitive shippers would likely pay more for better service quality; thus, guarantees for transport time, on-time arrival, and safety services could be offered to attract those shippers. These differentiated freight services, based on shippers' segmentation, will enhance shippers' perceptions and attitudes toward freight services and simultaneously increase railway companies' share in the freight market.

This paper focused on rail freight transport to study shippers' demand characteristics in relation to various rail services to determine the best interest of those shippers as they made their choices. This research area is important to railway companies in China to keep their market shares and will be further studied in the authors' research as it compares and analyzes the preference heterogeneity of multimodal shippers. In addition, the authors would like (*a*) to study possible impacts of preference heterogeneity on policy making, especially in relation to shippers' reactions to service changes in the real freight market, on the basis of more empirical studies in China and (*b*) to compare the preferences of different actors (shippers, consignees, logistics managers, etc.) in the supply chain. Comparative studies should also be performed between China and Europe to uncover differences with respect to contextual factors.

ACKNOWLEDGMENTS

The authors thank Maarten Kroesen, three anonymous referees, and the editor for their kind advice on this paper. The research was partly performed at Delft University of Technology and financially supported by the Southwest Jiaotong University Innovation Fund for PhD students (2014) as well as the State Scholarship Fund from the China Scholarship Council (2013). The fourth author acknowledges a grant from the Netherlands Organization for Scientific Research.

REFERENCES

1. *National Data: Transport, Postal and Telecommunication Services, 2005–2014.* National Bureau of Statistics of China. http://data.stats.gov.cn/easyquery.htm?cn=C01. Accessed July 9, 2015.
2. Crainic, T. G. Service Network Design in Freight Transportation. *European Journal of Operational Research,* Vol. 122, No. 2, 2000, pp. 272–288.
3. Shainesh, G., and M. Mathur. Service Quality Measurement: The Case of Railway Freight Services. *Vikalpa,* Vol. 25, No. 2, 2000, pp. 15–22.
4. Barnhart, C., N. Krishnan, D. Kim, and K. Ware. Network Design for Express Shipment Delivery. *Computational Optimization and Applications,* Vol. 21, No. 3, 2002, pp. 239–262.
5. Liang, G. S., T. Y. Chou, and S. F. Kan. Applying Fuzzy Quality Function Deployment to Identify Service Management Requirements for an Ocean Freight Forwarder. *Total Quality Management and Business Excellence,* Vol. 17, No. 5, 2006, pp. 539–554.
6. Wang, R. T. Improving Service Quality Using Quality Function Deployment: The Air Cargo Sector of China Airlines. *Journal of Air Transport Management,* Vol. 13, No. 4, 2007, pp. 221–228.
7. McGinnis, M. A. Shipper Attitudes Toward Freight Transportation Choice: A Factor Analytic Study. *International Journal of Physical Distribution and Materials Management,* Vol. 10, No. 1, 1979, pp. 25–34.
8. Matear, S., and R. Gray. Factors Influencing Freight Service Choice for Shippers and Freight Suppliers. *International Journal of Physical Distribution and Logistics Management,* Vol. 23, No. 2, 1993, pp. 25–35.
9. Cullinane, K., and N. Toy. Identifying Influential Attributes in Freight Route/Mode Choice Decisions: A Content Analysis. *Transportation Research Part E: Logistics and Transportation Review,* Vol. 36, No. 1, 2000, pp. 41–53.
10. Danielis, R., E. Marcucci, and L. Rotaris. Logistics Managers' Stated Preferences for Freight Service Attributes. *Transportation Research Part E: Logistics and Transportation Review,* Vol. 41, No. 3, 2005, pp. 201–215.
11. Fowkes, A. S., C. A. Nash, and G. Tweddle. Investigating the Market for Inter-Modal Freight Technologies. *Transportation Research Part A: Policy and Practice,* Vol. 25, No. 4, 1991, pp. 161–172.
12. Chu, H. C. Exploring Preference Heterogeneity of Air Freight Forwarders in the Choices of Carriers and Routes. *Journal of Air Transport Management,* Vol. 37, 2014, pp. 45–52.
13. Arunotayanun, K., and J. W. Polak. Taste Heterogeneity and Market Segmentation in Freight Shippers' Mode Choice Behaviour. *Transportation Research Part E: Logistics and Transportation Review,* Vol. 47, No. 2, 2011, pp. 138–148.
14. De Jong, G., M. Kouwenhoven, J. Bates, P. Koster, E. Verhoef, L. Tavasszy, and P. Warffemius. New SP Values of Time and Reliability for Freight Transport in the Netherlands. *Transportation Research Part E: Logistics and Transportation Review,* Vol. 64, 2014, pp. 71–87.
15. Fowkes, A. S., P. E. Firmin, A. E. Whiteing, and G. Tweddle. Freight Road User Valuations of Three Different Aspects Of Delay. *Proceedings of the AET European Transport Conference, Sept. 10–12, 2001,* Homerton College, Cambridge, United Kingdom.
16. Fowkes, A. S., P. E. Firmin, G. Tweddle, and A. E. Whiteing. How Highly Does the Freight Transport Industry Value Journey Time Reliability—and For What Reasons? *International Journal of Logistics: Research and Applications,* Vol. 7, No. 1, 2004, pp. 33–43.
17. Fowkes, T. The Design and Interpretation of Freight Stated Preference Experiments Seeking to Elicit Behavioural Valuations of Journey Attributes. *Transportation Research Part B: Methodological,* Vol. 41, No. 9, 2007, pp. 966–980.
18. Halse, A. H., and M. Killi. *Values of Transport Time and Reliability for Railway Freight.* Norwegian Institute of Transport Economics, Oslo, Norway, 2012.
19. Oum, T. H. A Cross Sectional Study of Freight Transport Demand and Rail–Truck Competition in Canada. *Bell Journal of Economics,* 1979, pp. 463–482.
20. Lu, C. S. The Impact of Carrier Service Attributes on Shipper–Carrier Partnering Relationships: A Shipper's Perspective. *Transportation Research Part E: Logistics and Transportation Review,* Vol. 39, No. 5, 2003, pp. 399–415.
21. Feo-Valero, M., L. García-Menéndez, and R. Garrido-Hidalgo. Valuing Freight Transport Time Using Transport Demand Modelling: A Bibliographical Review. *Transport Reviews.* Vol. 31, No. 5, 2011, pp. 625–651.
22. Tavasszy, L., and G. De Jong. *Modelling Freight Transport.* Elsevier, Amsterdam, Netherlands, 2013.
23. Train, K., and W. W. Wilson. Estimation on Stated-Preference Experiments Constructed from Revealed-Preference Choices. *Transportation Research Part B: Methodological,* Vol. 42, No. 3, 2008, pp. 191–203.
24. Patterson, Z. R., G. D. Ewing, and M. Haider. Shipper Preferences Suggest Strong Mistrust of Rail: Results from Stated Preference Carrier Choice Survey for Quebec City–Windsor Corridor in Canada. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2008,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 67–74.
25. Vieira, L. F. M. *The Value of Service in Freight Transportation.* MIT Press, Cambridge, Mass., 1992.
26. De Jong, G. Value of Freight Travel-Time Savings. *Handbooks in Transport, Volume 1: Handbook of Transport Modelling,* Elsevier, Oxford, United Kingdom, 2000, pp. 553–564.
27. Fosgerau, M. Investigating the Distribution of the Value of Travel Time Savings. *Transportation Research Part B: Methodological,* Vol. 40, No. 8, 2006, pp. 688–707.
28. De Jong, G., M. Kouwenhoven, E. Kroes, P. Rietveld, and P. Warffemius, Preliminary Monetary Values for the Reliability of Travel Times in Freight Transport. *European Journal of Transport and Infrastructure Research,* Vol. 9, No. 2, 2009, pp. 83–99.
29. Kurri, J., A. Sirkiä, and J. Mikola. Value of Time in Freight Transport in Finland. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1725,* TRB, National Research Council, Washington, D.C., 2000, pp. 26–30.
30. Danielis, R., and E. Marcucci. Attribute Cut-Offs in Freight Service Selection. *Transportation Research Part E: Logistics and Transportation Review,* Vol. 43, No. 5, 2007, pp. 506–515.

31. Bergantino, A. S., and S. Bolis. Monetary Values of Transport Service Attributes: Land Versus Maritime Ro-Ro Transport. An Application Using Adaptive Stated Preferences. *Maritime Policy and Management,* Vol. 35, No. 2, 2008, pp. 159–174.
32. Zamparini, L., J. Layaa, and W. Dullaert. Monetary Values of Freight Transport Quality Attributes: A Sample of Tanzanian Firms. *Journal of Transport Geography,* Vol. 19, No. 6, 2011, pp. 1222–1234.
33. Zamparini, L., and A. Reggiani. Freight Transport and the Value of Travel Time Savings: A Meta-Analysis of Empirical Studies. *Transport Reviews,* Vol. 27, No. 5, 2007, pp. 621–636.
34. Carson, R. T., and J. J. Louviere A Common Nomenclature for Stated Preference Elicitation Approaches. *Environmental and Resource Economics,* Vol. 49, No. 4, 2011, pp. 539–559.
35. Halme, M., and M. Kallio. Estimation Methods for Choice-Based Conjoint Analysis of Consumer Preferences. *European Journal of Operational Research,* Vol. 214, No. 1, 2011, pp. 160–167.
36. Rao, V. R. Choice Based Conjoint Studies: Design and Analysis. In *Applied Conjoint Analysis.* Springer, Berlin, 2014, pp. 127–183.
37. Ben-Akiva, M. E., and S. R. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand.* Vol. 9. MIT Press, Cambridge, Mass., 1985.
38. Chorus, C. G. A New Model Of Random Regret Minimization. *European Journal of Transport and Infrastructure Research,* Vol. 10, No. 2, 2010, pp. 181–196.
39. Chorus, C. G., S. van Cranenburgh, and T. Dekker. Random Regret Minimization for Consumer Choice Research: Assessment of Empirical Evidence. *Journal of Business Research,* Vol. 67, 2014, pp. 2428–2436.
40. Significance Quantitative Research, VU University Amsterdam, and John Bates Services. *Values of Time and Reliability in Passenger and Freight Transport in The Netherlands.* Ministry of Infrastructure and the Environment, The Hague, Netherlands, 2012.
41. Carrion, C., and D. Levinson. Value of Travel Time Reliability: A Review of Current Evidence. *Transportation Research Part A: Policy and Practice,* Vol. 46, No. 4, 2012, pp. 720–741.
42. Lam, T. C., and K. A. Small. The Value of Time and Reliability: Measurement from a Value Pricing Experiment. *Transportation Research Part E: Logistics and Transportation Review,* Vol. 37, No. 2, 2001, pp. 231–251.
43. Hollander, Y. Direct Versus Indirect Models for the Effects of Unreliability. *Transportation Research Part A: Policy and Practice,* Vol. 40, No. 9, 2006, pp. 699–711.
44. Greene, W. H. *NLOGIT 5 Reference Guide.* Econometric Software, Inc., Plainview, N.Y., 2012.
45. Vermunt, J. K., and J. Magidson. *Upgrade Manual for Latent GOLD Choice 5.0: Basic, Advanced, And Syntax.* Statistical Innovations, Inc., Belmont, Mass., 2013.
46. Lusheng, A. *New Practice of Railway Transportation in China.* China Railway Publishing House, Beijing, 2009.
47. Greene, W. H., and D. A. Hensher. A Latent Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit. *Transportation Research Part B: Methodological,* Vol. 37, No. 8, 2003, pp. 681–698.
48. Boeri, M., R. Scarpa, and C. G. Chorus. Stated Choices and Benefit Estimates in the Context of Traffic Calming Schemes: Utility Maximization, Regret Minimization, or Both? *Transportation Research Part A: Policy and Practice,* Vol. 61, 2014, pp. 121–135.
49. Hess, S., and C. G. Chorus. Utility Maximization and Regret Minimization: A Mixture of a Generalization. In *Bounded Rational Choice Behavior: Applications in Transport* (S. Rasouli and H. Timmermans, eds.), Emerald, Bingley, United Kingdom, 2015, pp. 31–50.
50. Hess, S., A. Stathopoulos, and A. Daly. Allowing for Heterogeneous Decision Rules in Discrete Choice Models: An Approach and Four Case Studies. *Transportation,* Vol. 39, No. 3, 2012, pp. 565–591.
51. Boeri, M., and L. Masiero. Regret Minimisation and Utility Maximisation in a Freight Transport Context. *Transportmetrica A: Transport Science,* Vol. 10, No. 6, 2014, pp. 548–560.

The Standing Committee on Freight Rail Transportation peer-reviewed this paper.