Modelling consumer perceptions of service quality for urban public transport systems using statistical models for ordinal data

Maria Iannario¹ · Anna Clara Monti²

Received: 15 February 2020 / Accepted: 6 January 2021 / Published online: 4 February 2021
© The Author(s) 2021

Abstract

The paper deals with ordinal response models to evaluate urban public transport systems with the purpose of interpreting consumers' responses with reference to their profiles. New methodological developments provide opportunities for a more thorough and accurate analysis of perceived service quality. The evaluation of the uncertainty component accounts for accuracy in the assessments. Diagnostic procedures allow to evaluate model specification, with respect to the proportional odds assumption, the adequacy of the mean structure and the occurrence of heterogeneity. The impact of the covariates on the discrete distribution of the observed response is appraised through their marginal effects. The selection of the appropriate covariates leads to the identification of clusters of users, which are compared through ordinal superiority measures. Consequently critical situations are detected.

Keywords Consumer perception · Marginal effects · Ordinal response models · Ordinal superiority measures · Transport system evaluation

1 Introduction

This contribution concerns statistical models for rating data aimed to interpret consumers' responses collected through a questionnaire on service evaluation of urban public transport systems.

The analysis of customer satisfaction related to service quality of transport companies is dealt with in several works for enhancing planning, quality management, schedule reliability and dispatching (see [7,8,26], among others). Studies concerning different means of transport elucidate that the concept of loyalty depends on users' intentions to continue using the service, their willingness to recommend it to others, and their overall satisfaction [31]. Understanding consumers' perceptions represents a fundamental step in developing policies for customer

✉ Maria Iannario
maria.iannario@unina.it

¹ Department of Political Sciences, University of Naples Federico II, Via L. Rodinó, 22, Naples, Italy

² Department of Law, Economics, Management and Quantitative Methods, University of Sannio, Benevento, Italy

satisfaction and retention and in implementing strategies for a competitive advantage in service delivery [9,21]. This understanding requires quantitative measurements based on data collected by means of surveys in which the assessments are expressed on ordinal scales.

The model extensively applied to analyse ratings is the cumulative model [17]. Recently an extension of this model in a two-component mixture model—the CUP—has been introduced by [29] to account for uncertainty in the response process. In this paper we consider both models to interpret data from a “transport service quality and customer satisfaction survey” carried out in 2012 in the urban area of Naples.

Aim of the contribution is investigating the most relevant features which drive the response and analysing marginal effects of the explanatory variables. Clusters of users are identified. Consequently ordinal superiority measures are applied for comparing the clusters’ rating distributions. The income, the means of transport and the availability of alternative systems of connection are the covariates that affect the evaluation of public transport.

The paper is organized as follows. The next section briefly reviews the statistical models useful for the analysis of ratings, while Sect. 3 introduces the data. The models, along with all the necessary diagnostics, are illustrated in Sect. 4. Section 5 analyses the marginal effects of the covariates, Sect. 6 derives ordinal superiority measures among clusters of commuters, whereas Sect. 7 outlines the probability distribution of the response for selected users’ profiles. Some final remarks conclude the paper.

2 Statistical models for consumers perception

Let $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)'$ be a random sample generated by an ordinal variable $Y \sim G(y)$ on the support $\{1, \dots, m\}$, where m is a known integer. We interpret Y_i as the rating expressed by the i -th subject on a definite item.

The two frameworks used for analysing rating data are derived by specifying $G(y)$. They are the cumulative model and the CUP model, respectively. The first is nested in the second one with the main difference given by the inclusion of the uncertainty in the response process.

2.1 The cumulative model

The cumulative model assumes that the ordinal response Y derives from the categorization of an underlying (continuous) latent variable Y^* . For any i -th statistical unit $Y_i = j$ when $\alpha_{j-1} < Y_i^* \leq \alpha_j$, where $-\infty = \alpha_0 < \alpha_1 < \dots < \alpha_{m-1} < \alpha_m = +\infty$ are the thresholds of the scale of Y^* . Let $\mathbf{X}_i = (X_{i1}, \dots, X_{ip})$ be the vector of covariates for the i -th subject. The dependence of the response on the covariates is established through the latent regression model

$$Y_i^* = \mathbf{X}_i \boldsymbol{\beta} + \epsilon_i, \quad i = 1, 2, \dots, n,$$

where $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)$ is the vector of regression coefficients, and $\epsilon_i \sim F_\epsilon(\cdot)$.

The probability mass function of Y_i , conditionally on $\mathbf{X}_i = \mathbf{x}_i = (x_{i1}, \dots, x_{ip})$, for $j = 1, \dots, m$, is

$$Pr(Y_i = j | \mathbf{x}_i) = Pr(\alpha_{j-1} < Y_i^* \leq \alpha_j | \mathbf{x}_i) = F_\epsilon(\alpha_j - \mathbf{x}_i \boldsymbol{\beta}) - F_\epsilon(\alpha_{j-1} - \mathbf{x}_i \boldsymbol{\beta}), \quad (1)$$

where $Pr(Y_i \leq j | \mathbf{x}_i) = F_\epsilon(\alpha_j - \mathbf{x}_i \boldsymbol{\beta})$ for $i = 1, \dots, n$ and $j = 1, 2, \dots, m$.

Common choices for $F_\epsilon(\cdot)$ are the Gaussian, the logistic, the (complementary) log-log and the Cauchy distribution, whose related models are named probit, logit, extreme value

and cumulative cauchit model, respectively. Here we focus on the logit link function with $F_e(u) = 1 / \{1 + \exp(-u)\}$.

A usual assumption in ordinal regression implies that the explanatory variables have the same effect on the cumulative logit regardless of the category of the response, i.e.

$$\text{logit} \{Pr(Y_i \leq j | X_i)\} = \log \left\{ \frac{Pr(Y_i \leq j | X_i)}{Pr(Y_i > j | X_i)} \right\} = \alpha_j - X_i \beta. \quad (2)$$

For given X_i , the logit is altered only by the intercepts α_j . Consequently, log-odds ratios for values of a single covariate X_{ik} —which differ by 1-unit—are constant. This is the *proportional odds assumption*, hence the related model is named Proportional Odds Model (POM). A test for this assumption is introduced in Sect. 4 (see also [4]).

Syntheses aimed to emphasise the effects of categorical and continuous covariates on ordinal responses have been proposed and summarized by many Authors, as in [2]. In this context, marginal effects (i.e., the change in the probability of the response due to a 1-unit change in one of the regressors) appear especially interesting (see [3] and reference therein).

Inference for this model is based on the likelihood function, whose expression can be found in [17]. Its maximization is carried out through iterative least squares estimation methods for Generalized Linear Models GLM (see [27]). The global validation of the fitted model is performed according to both likelihood-based methods and descriptive measures [33].

Further diagnostics can be based on various type of residuals, such as those recently introduced by [14].

2.2 The CUP model

The alternative modelling framework has been recently proposed by [29] who extend the class of cumulative models by introducing a further component to account for uncertainty in the response process. It mimics the idea of the CUB model [6,25] assuming that the response of each individual is the combination of a feeling attitude towards the item (preference part) and an intrinsic uncertainty component surrounding the discrete choice. The feeling attitude accounts for the subjective perception towards the object being evaluated and considers any reasoned assessment, as well as the set of emotions, awareness and sensitivity logically connected with the object. The uncertainty component accounts for indecision in the process of selecting the ordinal response and considers elements not logically connected with the item being evaluated, such as the unconscious willingness to delight the interviewer, the difficulty in evaluation using limited information, lack of self-confidence, laziness, boredom, etc. Notice that the uncertainty component is different from randomness which is inherent to sampling variability of surveys. The mixture has been denoted as CUP model, that is a Combination of a discrete Uniform random variable U and a Preference random variable Y^P . Formally, let Y_i be the response of an individual related to the explanatory variables X_i , the distribution of the CUP model for $j = 1, 2, \dots, m$ is

$$Pr(Y_i = j | X_i) = \pi Pr_M(Y_i^P = j | X_i) + (1 - \pi) Pr(U_i = j), \quad (3)$$

where Y_i^P and U_i are latent variables taking values from $1, \dots, m$. The distribution of Y_i^P is determined by $Pr_M(Y_i^P = j | X_i)$, which can be any ordinal model M , whereas $Pr(U_i = j) = 1/m$ represents the Uniform distribution. Here we consider model (1) for the probability of Y^P .

Table 1 Summary statistics concerning the characteristics of the 311 respondents of the survey

Gender

Female (55.63%)–Male (44.37)%

Education

High school diploma (96.15%), Bachelors degree or higher (3.85%)

Income (Household income sufficient to see the family

through to the end of the month—on a Likert scale from 1 (with great difficulty) to 10 (very easily))

1 (8.04%), 2 (6.75%), 3 (15.11%), 4 (22.19%), 5 (19.29%)

6 (14.50%), 7 (7.07%), 8 (3.20%), 9 (2.25%), 10 (1.60%)

Working condition

Employed (42.77%) Unemployed (57.23%)

Age

Min (15)–Max (47)

Equation (3) saves the traditional role of the predictors (as in the cumulative model) but considers the additional parameter π (as in the family of the CUB model) to assess the uncertainty component.

An extension of this model allows the uncertainty parameter to depend on a vector of subjects' covariates $\mathbf{Z}_i$, through a logit function, i.e. $\pi_i = 1 / \{1 + \exp(-\mathbf{Z}_i \boldsymbol{\eta})\}$. A further development replaces the Uniform distribution by a more flexible one to take into account different response styles [30].

Inference for this model is based on the likelihood function (see [29], for details).

3 Service quality and customer satisfaction in public transport of the urban area of Naples

The survey is an observational study on the service quality and customer satisfaction in public transport of people living in the urban area of Naples, Italy. It has been conceived and carried out by the Department of Political Sciences, University of Naples Federico II in December 2012. A large number of students has been randomly selected and each of them has been invited to recruit further subjects from their acquaintances by means of a *chain sampling*. The final sample is composed by 311 individuals, who were asked to rate their global perception of public transport, answering the question “Overall the transport system in the urban area of Naples is satisfactory?”, with responses on a 1–5 scale (1 = very unsatisfied; 5 = very satisfied). Further information on the individual characteristics of the subjects were collected along with the ratings. Table 1 reports their main features.

The frequency distribution of the global perception is illustrated in the left panel of Fig. 1, while the distribution of the mean of transport selected by respondents as the usual one is in the central panel. The dependence between the ratings and the means of transport is clearly evident in the mosaic plot displayed in the right panel of Fig. 1.

4 The model for the evaluation of the transport system

Initially a CUP model is fitted to the data. Here we assume that $Y_i^p = j$ when $\alpha_{j-1} < Y_i^* \leq \alpha_j$, with latent regression for the preference part given by

Fig. 1 Left: frequency distribution of the overall satisfaction. Center: frequency distribution of the means of transport. Right: mosaic plot for overall satisfaction and means of transport. In labels: sbus stands for *Suburban bus*, tube for *Tube*, boat for *Ferry boat*, train for *Train*, cbus for *Citybus*, and ltrain for *Local Train*

$$\begin{aligned}
 Y_i^* = & \beta_1 \text{Income}_i + \beta_2 \text{Alternative}_i + \gamma_1 \text{City bus}_i \\
 & + \gamma_2 \text{Ferry boat}_i + \gamma_3 \text{Train}_i + \gamma_4 \text{Tube}_i \\
 & + \gamma_5 \text{Suburban bus}_i + \epsilon_i, \quad i = 1, \dots, n.
 \end{aligned}$$

Income is a variable collected on a 10 points scale ranging from 1 (end of month very difficult to reach) to 10 (end of month very easy to reach), which we handle as numeric by referring to [1,5,10]. It represents a proxy of the household income. *Alternative* is a binary variable which takes value 1 if alternative means of transport are available for the interviewed person and zero otherwise. The local train (*Local Train*), known as “circumvesuviana”, is the baseline mean of transport. Consequently *CityBus*, *FerryBoat*, *Train*, *Tube* and *SuburbanBus* are binary variables which identify the usual mean of transport (when different from the local train). The parameters related to the means of transport are denoted by γ_r , for $r = 1, \dots, 5$, to emphasize that the covariates correspond to different categories of a single polytomous variable. They measure the impact on the preference component of the different means of transport with respect to the baseline mean (*Local Train*).

The fitted model is

$$\begin{aligned}
 Y_i^* = & 0.254 \text{Income}_i + 0.991 \text{Alternative}_i + 1.130 \text{City bus}_i \\
 & (0.097) \quad (0.327) \quad (0.412) \\
 & + 1.915 \text{Ferry boat}_i + 1.675 \text{Train}_i + 1.618 \text{Tube}_i \\
 & (0.702) \quad (0.480) \quad (0.396) \\
 & + 1.711 \text{Suburban bus}_i + \hat{\epsilon}_i, \quad i = 1, \dots, n, \\
 & (0.559)
 \end{aligned}$$

with estimated thresholds $\hat{\alpha}_1 = 1.551$, $\hat{\alpha}_2 = 2.836$, $\hat{\alpha}_3 = 4.432$, $\hat{\alpha}_4 = 5.777$. The estimate of the uncertainty parameter $\hat{\pi} = 0.99$ is a boundary value, and indicates that there is no uncertainty in the response process. The ratings are entirely reasoned assessments.

Consequently subsequent analyses are based on the POM. It yields

$$\begin{aligned}
 Y_i^* = & 0.259 \text{Income}_i + 1.035 \text{Alternative}_i + 1.120 \text{City bus}_i \\
 & (0.055) \quad (0.218) \quad (0.347) \\
 & + 1.904 \text{Ferry boat}_i + 1.672 \text{Train}_i + 1.606 \text{Tube}_i \\
 & (0.705) \quad (0.327) \quad (0.313) \\
 & + 1.714 \text{Suburan bus}_i + \hat{\epsilon}_i, \quad i = 1, \dots, n, \\
 & (0.434)
 \end{aligned}$$

Table 2 95% confidence intervals for the regression coefficients related to the means of transport

City Bus	Ferry Boat	Train	Tube	Suburban bus
(0.443, 1.806)	(0.523, 3.310)	(1.038, 2.321)	(0.998, 2.227)	(0.866, 2.573)

Table 3 Estimated model with different link function

Link function	Probit	Logit	Cauchy	Complementary log-log
AIC	871.86	869.86	883.91	878.14

with estimated thresholds $\hat{\alpha}_1 = 1.561$, $\hat{\alpha}_2 = 2.869$, $\hat{\alpha}_3 = 4.519$, $\hat{\alpha}_4 = 5.898$. The log-likelihood takes value -426.0395 .

Table 2 reports the 95% confidence intervals for the regression coefficients related to the means of transport. They largely overlap suggesting that the value of the parameter may be unique. Hence the null hypothesis $\gamma_1 = \dots = \gamma_5$ has been tested. The likelihood ratio statistic is 3.778 and the p -value of the test is 0.437, so that the null hypothesis is not rejected. There is no significant difference in the satisfaction of the commuters by the various means of transport. The only exception is given by the users of local train, whose ratings tend to be lower, as made evident by the positive $\hat{\gamma}$'s.

Consequently the attention focuses on the model

$$Y_i^* = \beta_1 \text{Income}_i + \beta_2 \text{Alternative}_i + \beta_3 \text{Local Train}_i + \epsilon_i$$

where *Local Train* is a binary variable which takes value 1 if the mean of transport is the local train and zero otherwise (any other mean of transport). The final model is

$$Y_i^* = \underset{(0.054)}{0.261} \text{Income}_i + \underset{(0.216)}{1.020} \text{Alternative}_i - \underset{(0.269)}{1.535} \text{Local Train}_i + \hat{\epsilon}_i \quad (4)$$

with estimated thresholds $\hat{\alpha}_1 = 0.034$, $\hat{\alpha}_2 = 1.335$, $\hat{\alpha}_3 = 2.966$ and $\hat{\alpha}_4 = 4.335$. The log-likelihood takes value -427.9287 .

When *Income* increases higher responses become more likely. Also the availability of alternative means of transport induces a more positive evaluation of the service. The commuters are more generous in rating the service when they are not frustrated by the absence of choice. Finally users of the local train are less satisfied than other commuters.

Although the logit link is often chosen for the straightforward interpretation of the results, in the present case this link is also the one which provides the best fit. Table 3 reports the value of the AIC for various link functions. The minimum is attained by the logit link.

When the proportional odds assumption fails to hold, for at least one covariate, the regression coefficients vary with the category j assumed by Y , and the parameter β in (1) and (2) is replaced by the parameter β_j corresponding to the response.

The proportional odds assumption may be formally tested via the Brant test [4]. The test can be carried out separately for each covariate or simultaneously for all the explanatory variables. The values of the test statistics and the corresponding p -values for each hypothesis are shown in Table 4: they confirm the validity of the *POM*.

The occurrence of heterogeneity effects can be tested by means of the cumulative-logit location-scale model of [17], which allows the scale σ of the latent variable distribution to

Table 4 Brant test for assessing parallel assumption

Test for	Wald Statistic	p-values
All covariates simultaneously	7.50	0.58
Income	5.37	0.15
Alternative	1.54	0.67
Local Train	0.98	0.81

Fig. 2 Surrogate residual diagnostic plots for *Income* (left panel), *Alternative* (central panel) and *Local Train* (right panel)

depend on the explanatory variables ($\sigma = \exp(X_i\delta)$). Equation (2) is modified as follows

$$\text{logit}(\Pr(Y_i \leq j|X_i)) = \frac{\alpha_j - X_i\beta}{\exp(X_i\delta)}$$

for $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m-1$. Absence of heterogeneity effects implies $\delta = 0$. The p -value of the test on this hypothesis is 0.366, so that no evidence of heterogeneity effects arises.

For checking misspecification of the latent mean structure we consider the surrogate residuals [14]. A continuous variable S with the same conditional distribution of the latent variable Y^* truncated on the interval (α_{j-1}, α_j) is generated when $Y = j$. The variable S acts as a surrogate for the ordinal response Y . The surrogate residual is defined as

$$R^S = S - E(S|X) = S - E(Y^*|X) = S - X\beta - \int_{-\infty}^{+\infty} \epsilon dF(\epsilon).$$

In practice, the residuals are computed by replacing the unknown parameters by their estimates. This yields the value of the i -th residual $r_i^S = s_i - X_i\hat{\beta} - \int_{-\infty}^{+\infty} \epsilon dF(\epsilon)$. Conditional sampling allows to consider the continuous space of the simulated data, rather than the discrete space of the original data. The plots of the surrogate residuals versus the covariates allow to check whether there is a misspecification in the latent mean structure. They are displayed in Fig. 2. The medians of the residuals are very close to zero confirming an adequate mean structure.

For model (4) the concordance index [12], that is a measure of predictive power that strictly uses ordinality, is 0.715 and R^2 of [19] is 0.245 whereas the pseudo R^2 of [18] is 0.089 indicating that the deviance for the model is 8.9% smaller than for the null model. Finally the

percentage of Correctly Classified Responses (*CCR*) is 0.383. It does not appreciably differ from that of more complex models.

5 Marginal effects

In the cumulative model the regression coefficients express the influence of the explanatory variables on the latent variable Y^* , rather than on the observable (discrete) response [15]. Marginal effects instead allow to evaluate how a change in the explanatory variables affects the probability of the response categories. Mood [20] and Landerman et al. [13] argue that marginal effects can be substantially more informative than odds ratios, especially for comparing groups that arise within the statistical units of the sample.

In case of a continuous explanatory variable X_k , for each category j of Y , the marginal effect is defined as

$$\frac{\partial P(Y = j|X = \mathbf{x}^*)}{\partial X_k}. \quad (5)$$

It measures the change in $P(Y = j|X = \mathbf{x}^*)$ for a 1-unit change in X_k evaluated at $X = \mathbf{x}^*$.

It is especially interesting to compute the marginal effects for the probability of the extreme responses, corresponding to an entirely unsatisfied ($Y = 1$) or a thoroughly satisfied subject ($Y = m$). In case of the logit link, [3] show that

$$\begin{aligned} \frac{\partial P(Y = 1|X = \mathbf{x}^*)}{\partial X_k} &= -\beta_k P(Y = 1|X = \mathbf{x}^*) \{1 - P(Y = 1|X = \mathbf{x}^*)\}, \\ \frac{\partial P(Y = m|X = \mathbf{x}^*)}{\partial X_k} &= \beta_k P(Y = m|X = \mathbf{x}^*) \{1 - P(Y = m|X = \mathbf{x}^*)\}. \end{aligned}$$

Their highest absolute value is $|\beta_k|/4$ which occurs when the probability of the category is $1/2$.

It is clear from (5) that the evaluation of the marginal effect varies with $\mathbf{x}^*$. Its magnitude depends on the value of the regressor X_k , whose effect is computed, as well as on the values of all the other explanatory variables in the model [15]. Various approaches have been suggested in literature to circumvent this problem (see [3] and [16, pag. 242–246]). The first considers the average marginal effect (*AME*) of X_k obtained as the mean of the marginal effects evaluated on the n sample values $\mathbf{x}_1, \dots, \mathbf{x}_n$. Alternatively, after computing the mean $\bar{\mathbf{x}}$ of $\mathbf{x}_1, \dots, \mathbf{x}_n$, the marginal effect at the mean (*MEM*) can be obtained by setting $\mathbf{x}^* = \bar{\mathbf{x}}$. Finally the marginal effect can be computed for representative values of the explanatory variables (*MER*). This last approach is recommended when some explanatory variables can be regarded as group variables which identify clusters of statistical units, as in the case of the binary variables *Alternative* and *Local Train*. In this circumstance it is of interest to compare the marginal effects separately for the different groups. Hence, the analysis applies the *MER* approach to compare the marginal effects between commuters who travel by the local train and those who use different means of transport, and between commuters who have alternative means of transport and those who have no choice. Representative values are considered also for *Income*, they are the minimum *Income* = 1, the mode *Income* = 4 and the maximum *Income* = 10.

The marginal effects of *Income* on the probability of $Y = 1$ and $Y = 5$ are shown in Table 5 for the four groups of commuters which arise from the binary variables. In parenthesis, below the estimated marginal effects, there are their standard errors computed by delta method. All

Table 5 Marginal effects for *income*

	<i>Income</i> = 1		<i>Income</i> = 4		<i>Income</i> = 10	
	Alternative No	Yes	Alternative No	Yes	Alternative No	Yes
<i>P</i> (<i>Y</i> = 1)						
Other transports	− 0.064 (0.015)	− 0.045 (0.014)	− 0.051 (0.012)	− 0.027 (0.007)	− 0.017 (0.003)	− 0.007 (0.001)
Local train	− 0.044 (0.008)	− 0.064 (0.012)	− 0.061 (0.013)	− 0.061 (0.014)	− 0.050 (0.007)	− 0.026 (0.006)
<i>P</i> (<i>Y</i> = 5)						
Other transports	0.004 (0.001)	0.011 (0.002)	0.009 (0.003)	0.022 (0.006)	0.034 (0.014)	0.058 (0.018)
Local train	0.001 (0.000)	0.003 (0.001)	0.002 (0.001)	0.006 (0.002)	0.009 (0.005)	0.023 (0.011)

Fig. 3 Marginal effects of *income* on *P*(*Y* = 1)

the estimated marginal effects are significantly different from zero. Their maximum absolute value is $\hat{\beta}_1/4 = 0.065$, and several marginal effects for the probability of $Y = 1$ are pretty close to this threshold especially for local train commuters and low incomes.

Increasing income reduces the probability of the extreme negative response ($Y = 1$) and consequently increases the probability of the extreme positive response ($Y = 5$). The impact of an increase of income is generally higher when the initial value is low and tapers off for higher incomes.

Figure 3 shows the marginal effects of *Income* on the probability of $Y = 1$ for the various combination of the binary variables. For each bar, the asymptotic 95% confidence interval is illustrated. It is computed as $ME \pm 1.96SE(ME)$ where ME is the marginal effect and $SE(ME)$ is its standard error.

Table 6 Marginal effects for *alternative* means of transport

	$P(Y = 1)$			$P(Y = 5)$		
	Income			Income		
	1	4	10	1	4	10
Other transports	-0.220 (0.049)	-0.151 (0.035)	-0.044 (0.016)	0.028 (0.010)	0.058 (0.016)	0.180 (0.049)
Local train	-0.216 (0.051)	-0.250 (0.051)	-0.148 (0.044)	0.006 (0.003)	0.014 (0.005)	0.059 (0.025)

Table 7 Marginal effects for *Local Train*

	$P(Y = 1)$			$P(Y = 5)$		
	Income			Income		
	1	4	10	1	4	10
Alternatives NO	0.344 (0.055)	0.361 (0.060)	0.190 (0.059)	- 0.013 (0.005)	- 0.028 (0.008)	- 0.114 (0.037)
Alternatives YES	0.348 (0.064)	0.263 (0.059)	0.086 (0.034)	- 0.035 (0.011)	-0.072 (0.017)	- 0.235 (0.056)

When X_k is a binary variable, its marginal effect is computed as

$$P(Y = j|X_{(k)} = \mathbf{x}_{(k)}^*, X_k = 1) - P(Y = j|X_{(k)} = \mathbf{x}_{(k)}^*, X_k = 0)$$

where $X_{(k)}$ is the vector of the explanatory variables after removing X_k .

The marginal effects of the availability of alternative means of transport with respect to the usual one are displayed in Table 6. The estimates of all the marginal effects are significantly different from zero. Their maximum absolute value is $\hat{\beta}_2/4 = 0.255$.

The availability of alternative means of transport makes the passengers less critical, and reduces their propensity towards the extreme negative response ($Y = 1$). This beneficial effect increases when they travel by the local train or have a low income. When the choice of the mean of transport is not a compulsory one, the commuters tend to be more satisfied.

The marginal effects of *Alternative* along with the asymptotic confidence intervals are displayed in the left panel of Fig. 4.

Table 7 shows the marginal effects for the local train. Also in this case all the estimates of the marginal effects are significantly different from zero. Their maximum absolute value is $\hat{\beta}_3/4 = 0.384$.

Traveling by the local train makes the commuters more critical and increases the probability of the lowest response ($Y = 1$). This effect becomes sharper when no alternatives are available and the income decreases (the marginal effect for the probability of $Y = 1$ is very close to their maximum).

The marginal effects of *Local Train* are illustrated in the right panel of Fig. 4.

6 Ordinal superiority measures among clusters of users

The two binary variables of model (4) identify groups of subjects according to whether they travel or not by the local train and whether they have or not an alternative mean of transport.

Fig. 4 Marginal effects of *Alternative* means of transport (left panel) and *Local Train* (right panel) on $P(Y = 1)$

An approach to compare groups analyses the probability that an observation from one group falls above an independent observation from an alternative group. To this purpose it is useful to consider the ordinal superiority measures.

Let Y_{g_1} and Y_{g_2} be independent ordinal responses from two groups of statistical units, denoted by g_1 and g_2 . Agresti and Kateri [2] define the ordinal superiority measures

$$\gamma = P(Y_{g_1} > Y_{g_2}) + \frac{1}{2}P(Y_{g_1} = Y_{g_2}),$$

and

$$\Delta = P(Y_{g_1} > Y_{g_2}) - P(Y_{g_2} > Y_{g_1}),$$

where

$$\gamma \in [0, 1], \quad \Delta \in [-1, 1], \quad (6)$$

and

$$\gamma = (\Delta + 1)/2, \quad \Delta = 2\gamma - 1. \quad (7)$$

When $\gamma > 1/2$ or $\Delta > 0$, the ratings from g_1 tend to be larger than the ratings from g_2 assessing an ordinal superiority of g_1 over g_2 .

Agresti and Kateri [2] prove that for the logit link γ can be approximated as follows

$$\gamma \approx \frac{\exp(\hat{\beta}/\sqrt{2})}{1 + \exp(\hat{\beta}/\sqrt{2})}, \quad (8)$$

where $\hat{\beta}$ in (8) is the maximum likelihood estimate of the group effect.

By considering the binary variables *Alternative* and *Local Train* as group variables, the estimates of γ and Δ are displayed in Table 8 along with the 95% confidence intervals. Consistently with the results of the previous sections, the value of $\hat{\gamma}_{\text{Alternative}} = 0.673$ is significantly greater than $1/2$. It indicates that the ratings of commuters who have an alternative mean of transport are higher than the ratings of people without choice with probability 0.673. On the contrary, the estimate $\hat{\gamma}_{\text{Local Train}} = 0.252$ is significantly smaller than $1/2$ and markedly low. This result points out that the ratings of the commuters by the local train are lower than the ratings by other commuters with probability 0.748, providing further evidence of the poor level of satisfaction associated to this mean of transport.

Table 8 Ordinal superiority measures between groups of commuters

	$\hat{\gamma}$	$CI(\gamma)$	$\hat{\Delta}$	$CI(\Delta)$
Alternatives	0.673	(0.604, 0.736)	0.346	(0.208, 0.472)
Local Train	0.252	(0.188, 0.328)	− 0.496	(− 0.624, − 0.344)

Table 9 Ordinal superiority measures between groups of commuters identified by the binary covariates (*Alternative* and *Local Train*) at the modal income

Group 1		Group 2		γ	$CI(\gamma)$	Δ	$CI(\Delta)$
Local		Local					
Alternative	Train	Alternative	Train				
Yes	No	No	No	0.653	(0.594, 0.713)	0.306	(0.187, 0.425)
Yes	Yes	No	Yes	0.640	(0.583, 0.697)	0.280	(0.165, 0.394)
Yes	No	No	Yes	0.830	(0.773, 0.888)	0.660	(0.545, 0.776)
No	No	Yes	Yes	0.578	(0.475, 0.682)	0.156	(− 0.050, 0.363)
No	No	No	Yes	0.712	(0.650, 0.774)	0.423	(0.300, 0.547)
Yes	No	Yes	Yes	0.723	(0.656, 0.789)	0.445	(0.312, 0.578)

Alternative ordinal superiority measures for comparing groups are the following [2]

$$\gamma^{\tau}(g_1, g_2, \tilde{\mathbf{x}}_0) = \sum_{j>k} \hat{\tau}_{g_1,j}(\tilde{\mathbf{x}}_0) \hat{\tau}_{g_2,k}(\tilde{\mathbf{x}}_0) + \frac{1}{2} \sum_j \hat{\tau}_{g_1,j}(\tilde{\mathbf{x}}_0) \hat{\tau}_{g_2,j}(\tilde{\mathbf{x}}_0).$$

and

$$\Delta^{\tau}(g_1, g_2, \tilde{\mathbf{x}}_0) = \sum_{j>k} \hat{\tau}_{g_1,j}(\tilde{\mathbf{x}}_0) \hat{\tau}_{g_2,k}(\tilde{\mathbf{x}}_0) - \sum_{k>j} \hat{\tau}_{g_1,j}(\tilde{\mathbf{x}}_0) \hat{\tau}_{g_2,k}(\tilde{\mathbf{x}}_0)$$

where

$$\hat{\tau}_{g_{\ell},j}(\tilde{\mathbf{x}}_0) = \hat{P}(Y = j; g = g_{\ell}, \tilde{\mathbf{x}}_0), \quad \ell = 1, 2, \quad j = 1, \dots, m$$

is the estimated probability of $Y = j$ for the group identified by g_{ℓ} while the other covariates are fixed at the selected value $\tilde{\mathbf{x}}_0$. Also in this case (6) and (7) hold.

The binary covariates (*Alternative* and *Local Train*) identify the following four groups of users

1. commuters by local train without alternative,
2. commuters by local train with alternative,
3. commuters by other means of transport without alternative,
4. commuters by other means of transport with alternative.

The probability distribution of Y for these groups has been estimated from model (4), with selected value for income fixed at the mode ($Income = 4$). It yields the ordinal superiority measures reported in Table 9.

With the exception of the comparison between users of other means of transport without alternatives and local train commuters with alternatives, $\hat{\gamma}$ ($\hat{\Delta}$) is significantly different from 1/2 (zero), pointing out that the users from *Group 1* (in the table) tend to express larger ratings than users from *Group 2*. The results confirm the ordinal superiority of the ratings of people who have alternatives with respect to those without choice, and of the ratings

Fig. 5 Ordinal superiority measures among groups of commuters

Fig. 6 Estimated probability distribution of the response for selected profiles. Gray lines are referred to users who usually use the *Local Train* whereas black dashed lines are referred to consumers who traditionally use *Other Transports*

of commuters by other means of transport with respect to local train users. Furthermore, it has to be noticed that the ratings of commuters who do not travel by the local train and have alternatives enjoy ordinal superiority with respect to all the other three groups. These outcomes are displayed in Fig. 5.

7 Probability distribution of the responses for selected users' profiles

By selecting the values of the covariates appearing in model (4), users' profiles depending on income, the usual mean of transport and the availability of other services can be identified.

Figure 6 shows the estimated probability distribution of the response for selected profiles. The distribution of local train users is systematically shifted leftward, concentrating larger probability on lower ratings. The availability of alternative means of transport instead moves

the probability from the lower categories to the higher ones. Finally, when income increases the probability of lower ratings decreases.

8 Concluding remarks

Aim of the paper has been to show, through real-life data on service quality of urban transport systems, how models for ordinal responses can be used in the study and interpretation of human judgments, whenever opinions are measured by ratings.

The main points are: an appropriate treatment of the ordinal nature of the data; the possibility to measure both the level of the latent variable and the uncertainty surrounding the selection of the rating; the assessment of the impact of the individual characteristics on the variable of interest, the identification of clusters of users and the detection of critical opinions. Furthermore marginal effects measure the impact of a change in the explanatory variables on the discrete probability distribution of the response, while ordinal superiority measures allow to investigate differences among clusters of consumers and identify critical situations. These methodological tools are useful any time ratings need to be analysed in the field of service quality evaluation as for instance when one of the most famous tool ‘SERVQUAL’ is provided [22–24].

With reference to transport data, results show that there is no uncertainty in the response process, i.e. the ratings are entirely reasoned assessments. No difference arises among different means of transport, with the exception of the local train. Travelling by the local train makes the commuters more critical, while the availability of alternative means of transport increases their propensity towards positive responses. Also an increase in income makes the interviewee more generous in evaluation. The two binary variables *Alternative* and *Local Train* identify four clusters of users. The ratings of commuters who do not travel by the local train and have alternative enjoy ordinal superiority (a comparatively higher evaluation of service) with respect to all the other groups of respondents. The cluster of the local train commuters with no alternative is the most critical one, identifying the mean of transport where corrective actions should be primarily focussed.

The present analysis provides a first insight; which can be extended by using more comprehensive questionnaire data including additional information such as the point of origin/destination route, the reason of travel, the frequency of use of the mean of transport, the time spent for travelling and so forth. Finally, an alternative approach to investigate commuters’ satisfaction is the analysis of multiple item questionnaires through the Item Response Theory, finalized to assess the ‘ability’ of users to be satisfied (see [11,28,32] for details). Useful models to this purpose are the partial credit model and the graded response model, among others.

Funding Open Access funding provided by Università degli Studi di Napoli Federico II.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Agresti, A.: Analysis of Ordinal Categorical Data, 2nd edn. Wiley, Hoboken (2010)
2. Agresti, A., Kateri, M.: Ordinal probability effect measures for group comparisons in multinomial cumulative link models. *Biometrics* **73**, 214–219 (2017)
3. Agresti, A., Tarantola, C.: Simple ways to interpret effects in modeling ordinal categorical data. *Stat. Neerlandica* **72**, 210–223 (2018)
4. Brant, R.: Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics* **46**, 1171–1178 (1990)
5. Capecchi, S., Iannario, M., Simone, R.: Well-being and relational goods: a model-based approach to detect significant relationships. *Soc. Indic. Res.* **135**, 729–750 (2018)
6. D'Elia, A., Piccolo, D.: A mixture model for preference data analysis. *Comput. Stat. Data Anal.* **49**, 917–934 (2005)
7. Eboli, L., Mazzulla, G.: Service quality attributes affecting customer satisfaction for bus transit. *J. Public Transp.* **10**, 2 (2007)
8. Fischer, J.M., Mshadoni, S., Kennedy, A.A.: Why and how to use customer opinions: a quality-of-life and customer satisfaction-oriented foundation for performance-based decision-making. *Transp. Rev.* **34**, 86–101 (2014)
9. Fu, X., Juan, Z.: Understanding public transit use behavior: integration of the theory of planned behavior and the customer satisfaction theory. *Transportation* **44**, 1021–1042 (2017)
10. Gambacorta, R., Iannario, M.: Measuring job satisfaction with CUB models. *Labour* **27**(2), 198–224 (2013)
11. Hambleton, R.K., Swaminathan, H.: Item Response Theory: Principles and Applications. Kluwer Nijhoff, Boston (1985)
12. Harrell, F.E., Lee, K.L., Mark, D.B.: Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* **15**(4), 361–387 (1996)
13. Landerman, L.R., Mustillo, S.A., Land, K.C.: Modeling repeated measures of dichotomous data: testing whether the within-person trajectory of change varies across levels of between-person factors. *Soc. Sci. Res.* **40**(5), 1456–1464 (2011)
14. Liu, D., Zhang, H.: Residuals and diagnostics for ordinal regression models: a surrogate approach. *J. Am. Stat. Assoc.* **113**, 845–854 (2018)
15. Long, J.S.: Regression Models for Categorical and Limited Dependent Variables. Sage Publications Inc, New York (1997)
16. Long, J., Freese, J.: Regression Models for Categorical Dependent Variables Using STATA. Stata Press, College Station (2014)
17. McCullagh, P.: Regression models for ordinal data (with discussion). *J. R. Stat. Soc. Ser. B* **42**, 109–142 (1980)
18. McFadden, K.: Modeling the choice of residential location. In: Karlqvist, A., et al. (eds.) *Spatial Interaction Theory and Residential Location*, pp. 75–76. North-Holland, Amsterdam (1978)
19. McKelvey, R.D., Zavoina, W.: A statistical model for the analysis of ordinal level dependent variables. *J. Math. Sociol.* **4**, 103–120 (1975)
20. Mood, C.: Logistic regression: why we cannot do what we think we can do, and what we can do about it. *Eur. Sociol. Rev.* **26**(1), 67–82 (2010)
21. Mortona, C., Caulfield, B., Anable, J.: Customer perceptions of quality of service in public transport: Evidence for bus transit in Scotland. *Case Studies on Transport Policy*, **4**, 199–207 (2016) Mood, C.: Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, **26**(1), 67–82 (2010)
22. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: A conceptual model of service quality and its implication. *J. Mark.* **49**, 41–50 (1985)
23. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: SERVQUAL: a multi-item scale for measuring consumer perceptions of the service quality. *J. Retail.* **64**, 12–40 (1988)
24. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: Refinement and reassessment of the SERVQUAL scale. *J. Retail.* **67**, 420–450 (1991)
25. Piccolo, D.: On the moments of a mixture of uniform and shifted binomial random variables. *Quad. Stat.* **5**, 85–104 (2003)
26. Stelzer, A., Englert, F., Hörold, S., Mayas, M.: Improving service quality in public transportation systems using automated customer feedback. *Transp. Res. Part E Logist Transp. Rev.* **89**, 259–271 (2016)
27. Tutz, G.: Regression for Categorical Data. Cambridge University Press, Cambridge (2012)

28. Tutz, G.: Hierarchical models for the analysis of Likert scales in regression and item response analysis. *Int. Stat. Rev.* (2020). <https://doi.org/10.1111/insr.12396>
29. Tutz, G., Schneider, M., Iannario, M., Piccolo, D.: Mixture models for ordinal responses to account for uncertainty of choice. *Adv. Data Anal. Classif.* **11**, 281–305 (2017)
30. Tutz, G., Schneider, M.: Mixture models for ordinal responses with a flexible uncertainty component. *J. Appl. Stat.* **46**, 1–20 (2018)
31. van Lierop, D., Badami, M.G., El-Geneidy, A.M.: What influences satisfaction and loyalty in public transport? A review of the literature. *Transp. Rev.* **38**, 52–72 (2018)
32. Van der Linden, W.J., Hambleton, R.K.: *Handbook of Modern Item Response Theory*. Springer, New York (1997)
33. Veall, M.R., Zimmermann, K.F.: Pseudo- R^2 measures for some common limited dependent variable models. *J. Econ. Surv.* **10**, 241–259 (1996)

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